

**Visualisierung von impliziten geographischen Informationen durch
kartenähnliche Graphiken**

Von der
Carl-Friedrich-Gauß-Fakultät
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades einer
Doktoringenieurin (Dr.-Ing.)

genehmigte Dissertation

von
Sarah Tauscher
geboren am 26.01.1981
in Hattingen

Eingereicht am:	28.03.2018
Disputation am:	26.10.2018
1. Referent:	PD Dr. Karl Neumann
2. Referentin:	Prof. Dr. Monika Sester

2018

ABSTRACT

A lot of user generated information accumulated in the web is related to a place, with the location usually being just one piece of information among many, which gets no special attention. However, more and more geographic information is collected by laymen and published on the web. Despite the technical possibilities that are offered in Web 2.0, it is usually not possible for users without expert knowledge to produce legible and appealing maps. This problem exists because the user determines the content to be displayed without checking whether the resulting map meets the cartographic requirements. Furthermore, the user has no possibilities to edit the standard map images, for example, hide for his subject irrelevant objects or reducing overlap of relevant objects by displacement. Especially when displaying POIs, overlaps between point symbols often occur. Despite the need, there is no established method for displacing point data for various reasons.

Therefore, the focus of this work is the development of methods for the displacement of point signatures. Voronoi diagrams are used as auxiliary structures and sentiment is visualized as user-generated information. For the design of the visualizations relevant cartographic requirements are taken into account and evaluated by quality measures. For the depiction of sentiments, in addition to the use of point symbols, two further types of visualizations are created: adaptation of given map symbols and the representation of sentiments as continua. Displacement techniques for point symbols are designed and implemented. To determine the direction of displacement two different heuristics are proposed and examined. Furthermore, a way to increase the efficiency by dividing the point set is shown. The evaluation of the designed point symbols is done by a survey. Subsequently, the realized method for circular symbols with equal size is evaluated in three aspects: degree of reduction of the iteration steps by decomposition of the point set, achieved reduction of the overlap area and change of the relative position of the points.

ZUSAMMENFASSUNG

Viele von Web 2.0-Benutzern gesammelte Daten sind ortsbezogen, wobei der Ort meist nur eine Information unter vielen ist und dem Ortsbezug häufig keine besondere Bedeutung zugesprochen wird. Jedoch werden auch mehr und mehr geographische Informationen von Laien gesammelt und im Netz veröffentlicht. Trotz der technischen Möglichkeiten, die im Web-2.0 geboten werden, ist es für Nutzer ohne entsprechendes Expertenwissen meist nicht möglich, gut lesbare und ansprechende Karten zu erzeugen. Dieses Problem besteht, da der Nutzer den darzustellenden Inhalt bestimmt, ohne dass überprüft wird, ob die daraus resultierende Karte den kartographischen Ansprüchen genügt. Des Weiteren hat der Nutzer keinerlei Möglichkeiten, die Standardkartenbilder zu bearbeiten, um beispielsweise für seine Thematik irrelevante Objekte auszublenden oder durch Verdrängung relevante Objekte freizustellen. Besonders bei der Darstellung von POIs treten Überlappungen zwischen Signaturen häufig auf. Trotz des Bedarfs existiert aus verschiedenen Gründen keine etablierte Methode zur Verdrängung von Punktdaten.

Daher liegt der Schwerpunkt dieser Arbeit auf der Entwicklung von Verfahren zur Verdrängung von Punktsignaturen. Als Hilfsstrukturen werden dazu Voronoi-Diagramme benutzt und als nutzergenerierte Information werden Sentiments visualisiert. Für den Entwurf der Visualisierungen werden relevante kartographische Bedingungen berücksichtigt und durch zugehörige Qualitätsmaße bewertet. Für die Darstellung von Sentiments werden neben der Verwendung von Punktsignaturen zwei weitere Darstellungsarten erstellt: Anpassung gegebener Signaturen und die Darstellung von Sentiments als Kontinua. Es werden Verdrängungsverfahren für Punktsignaturen entworfen und implementiert. Zur Bestimmung der Verschiebungsrichtung werden zwei verschiedenen Heuristiken vorgeschlagen und untersucht. Des Weiteren wird eine Möglichkeit zur Steigerung der Effizienz durch Aufteilung der Punktmenge aufgezeigt. Die Bewertung der entworfenen Punktsignaturen erfolgt durch eine Umfrage. Anschließend wird das realisierte Verfahren für gleich große Kreissignaturen in drei Aspekten evaluiert: Grad der Reduzierung der Iterationsschritte durch Zerlegung der Punktmenge, erreichte Verminderung der Überlappungsfläche und Veränderung der relativen Lage der Punkte.

Danksagung

An erster Stelle möchte ich mich ganz herzlich bei PD Dr. Karl Neumann bedanken, der mir über Jahre stets geduldig und beratend zur Seite gestanden hat. Des Weiteren danke ich Prof. Dr. Wolf-Tilo Balke für die Zeit am Institut und die angenehme Arbeitsatmosphäre. Mein besonderer Dank gilt Prof. Dr. Monika Sester für die inspirierenden fachlichen Diskussionen und interessierten Nachfragen.

Von ganzem Herzen möchte ich auch meiner Familie und meinem Freund danken, die mich immer unterstützt haben.

Inhaltsverzeichnis

EINLEITUNG	3
1.1. ZIEL DIESER ARBEIT	6
1.2. AUFBAU DER ARBEIT	7
VORONOI-DIAGRAMME.....	9
2.1. EIGENSCHAFTEN.....	10
2.2. SCHWERPUNKT-VORONOI-DIAGRAMM	10
2.3. POWER-DIAGRAMME	12
GRUNDLAGEN DER INFORMATIONSVISUALISIERUNG	15
3.1. CHARAKTERISIERUNG DER DARZUSTELLENDEN INFORMATION	16
3.2. GRAPHISCHE VARIABLEN.....	17
3.3. WAHRNEHMUNGSARTEN VON GRAPHISCHEN VARIABLEN.....	20
SENTIMENTS	23
4.1. PSYCHOLOGISCHE MODELLE	23
4.2. EXTRAKTION AUS TEXTEN.....	24
4.2.1. Sentiment-Analyse von räumlichen Daten	25
4.2.2. Wörterbuch-basierte Verfahren.	26
4.3. VISUALISIERUNG VON SENTIMENTS	26
4.4. KARTOGRAPHISCHE DARSTELLUNG.....	29
GENERALISIERUNG VON POI-KARTEN.....	33
5.1. VERDRÄNGUNG	34
5.2. KARTOGRAPHISCHE BEDINGUNGEN UND QUALITÄTSMASSE	36
5.3. LÖSBARKEIT VON ÜBERLAPPUNGEN.....	38
ENTWURF VON SIGNATUREN FÜR SENTIMENTS.....	43
6.1. KOMMUNIKATIONSZIEL	44
6.2. DARSTELLUNG VON SENTIMENTS DURCH ANPASSUNG VON SIGNATUREN	46
6.3. DARSTELLUNG VON SENTIMENTS ALS KONTINUA	50
6.4. PUNKTSIGNATUREN FÜR KLEINMAßSTÄBIGE KARTEN.....	53
6.5. FAZIT	55
VERDRÄNGUNGSVERFAHREN FÜR PUNKTSIGNATUREN	57
7.1. VERDRÄNGUNG GLEICH GROßER KREISFÖRMIGER SIGNATUREN	58
7.1.1 Heuristiken zur Bestimmung der Verschiebungsrichtung	59
7.1.2 Beschränkung der Fläche der Voronoizellen	64
7.1.3 Zerlegung in unabhängige Teilmengen.....	66
7.1.4 Auswahl der Parameter	69
7.2. VERALLGEMEINERUNG DES VERFAHRENS	71

7.2.1.	Symbole unterschiedlicher Größe	72
7.2.2.	Lokale Änderung der Zellgeometrie zur Berechnung der Verschiebungsrichtung	73
7.2.3.	Verdrängung elliptischer Symbole.....	75
RESULTATE DER EXPERIMENTE		77
8.1.	BENUTZERSTUDIE ZUR LESBARKEIT DER PUNKTSIGNATUREN	77
8.2.	EVALUATION DES VERDRÄNGUNGSVERFAHRENS.....	80
8.2.1.	Testdatensätze.....	80
8.2.2.	Zerlegung der Punktmenge	82
8.2.3.	Selektion zur Vermeidung unlösbarer Überlappungen	84
8.2.4.	Lesbarkeit der Signaturen	86
8.2.5.	Erhaltung der räumlichen Anordnung.....	87
8.2.6.	Fazit.....	90
8.3.	VERGLEICH ZWISCHEN VERWENDUNG DES POWER- UND DES VORONOI- DIAGRAMMS	94
8.4.	ANPASSEN DER ZELLGEOMETRIE.....	97
8.5.	ABSCHÄTZUNG DES SIGNATURRADIUS.....	98
ZUSAMMENFASSUNG UND AUSBLICK		103
ABBILDUNGSVERZEICHNIS.....		107
LITERATURVERZEICHNIS		111

Einleitung

Der Erfolg typischer Web 2.0 Anwendungen wie Facebook, Twitter, Flickr, Youtube verdeutlicht, dass der soziale Aspekt des World Wide Webs mehr und mehr an Bedeutung gewinnt. Dank dieser Services werden kontinuierlich große Datenmengen generiert, die von schlichten "Likes" über Kurznachrichten, Kommentaren und Blogs bis hin zu Fotos und Videos reichen. Viele dieser nutzergenerierten Daten sind ortsbezogen, wobei der Ort meist nur eine Information unter vielen ist und dem Ortsbezug keine besondere Bedeutung zugesprochen wird. Jedoch werden auch mehr und mehr geographische Informationen von Laien gesammelt und im Netz veröffentlicht [Go07]. Räumliche Daten, die nach den Prinzipien der nutzergetriebenen Inhalte des Web 2.0 erfasst werden, werden als Volunteered Geographic Information (VGI) bezeichnet. Ein Beispiel hierfür ist Wikimapia, das Benutzern erlaubt, Informationen über interessante Orte bereitzustellen. Jeder Eintrag in Wikimapia besteht aus einer textuellen Beschreibung und einem georeferenzierten Rechteck. Die Darstellung erfolgt als Punkt auf einem Satellitenbild, über den der Zugriff auf die Beschreibung erfolgt. Ein weiteres bekanntes Beispiel ist Flickr, das mehr als 20 Millionen mit GPS-Koordinaten versehenen Fotos bereitstellt. Eine Möglichkeit, räumliche Daten differenzierter zu erfassen und abzurufen, bietet das OpenStreetMap Projekt, für dessen Nutzung jedoch ein gewisses Expertenwissen notwendig ist. Die Mitwirkenden erstellen entweder eigene GPS-Tracks oder sammeln andere frei verfügbare Geodaten sowie weitere Metadaten als Tags. All diese Daten werden kombiniert und als Karte sowie als Rohdaten im Netz veröffentlicht. Eine weitere sehr populäre Art, um räumliche Daten zu teilen, sind einfach zu benutzende Web 2.0 Umgebungen, wie Google Maps. Diese Services ermöglichen die Generierung von Karten mit Hilfe von Application Programming Interfaces (API). Ein Beispiel ist Google My Maps, das 2007 veröffentlicht wurde. Es bietet die Möglichkeit Karten mit benutzerdefinierten Inhalten über ein Interface zu erstellen, dessen Bedienung keinerlei kartographische Kenntnisse benötigt [SW12]. Hierzu werden Layer mit beliebigem Inhalt erzeugt, die anschließend über die vorhandenen Karten- und Satellitenbilder von Google Maps visualisiert werden können. Nachteile dieser Services sind, dass nur wenige Hintergrundkarten zur Verfügung stehen und die Möglichkeiten zur Anpassung der Hintergrundkarten an die von Benutzer hinzugefügten Daten stark limitiert sind. Des Weiteren wird weder eine Hilfestellung zur Einhaltung kartographischer Bedingungen bei der Generierung einer Karte noch die Möglichkeit einer automatischen Qualitätskontrolle der fertigen Karte zur Verfügung gestellt. Eine zweite weit verbreitete Option Karten über das Internet zur Verfügung zu stellen, ist die Verwendung von Web Map Services (WMSs) [OGC02] und weiteren offenen Standards, die vom Open Geo-

4 Einleitung

Spatial Consortium (OGC) definiert wurden. Im Hinblick auf die kartographische Qualität bestehen bei den meisten WMSs ebenfalls Defizite, allerdings können diese durch kartographische Erweiterungen des Symbology Encoding Standards eliminiert werden [IHH10]. Des Weiteren ist bei der Verwendung vom WMS zu beachten, dass, obwohl Kartenlayer von verschiedenen Anbietern einfach kombiniert werden können, d.h. die Interoperabilität garantiert ist, die resultierende Karte oft nicht hilfreich und/oder schlecht lesbar ist, falls die Layer nicht zur Kombination vorgesehen sind [WA02].

Trotz all dieser technischen Möglichkeiten zur Kartenerzeugung ist es für Nutzer ohne entsprechendes Expertenwissen meist nicht möglich, gut lesbare und ansprechende Karten zu erzeugen. Wenn nur einige wenige Orte oder eine Route in großem Maßstab dargestellt werden soll, kann das Ergebnis durchaus ansprechend sein, wie in Abbildung 1 an der mit Hilfe von Google Maps erstellten Karte zu sehen ist, die den kürzesten Weg vom Braunschweiger Hauptbahnhof über das Schloss und die Burg Dankwarderode zum Altstadtmarkt visualisiert.

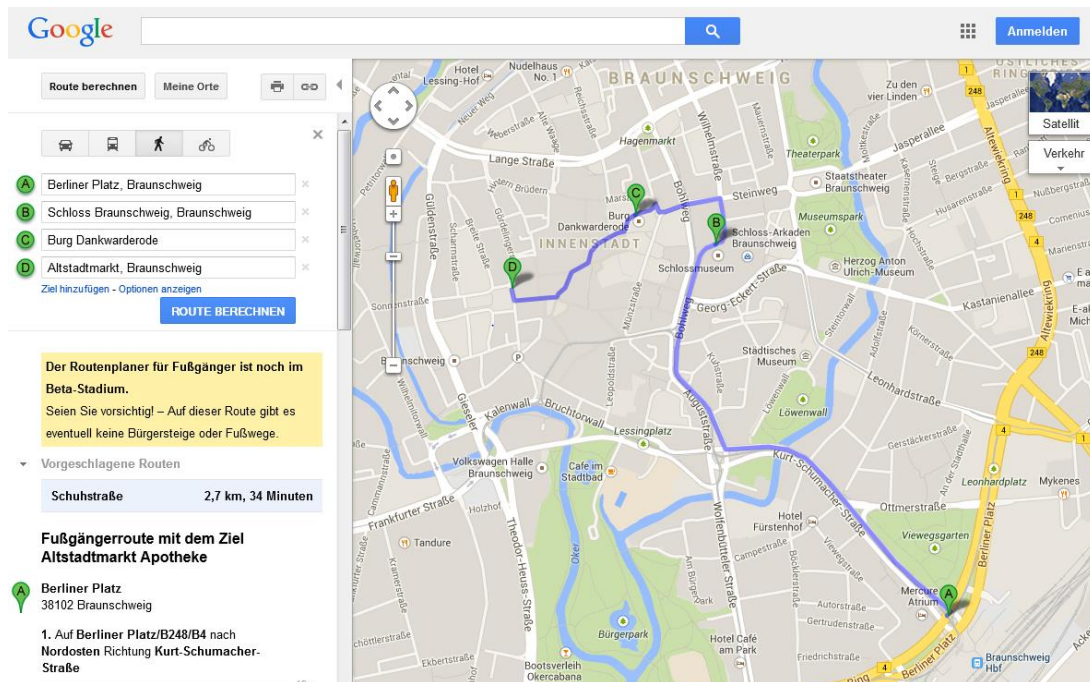


Abb. 1: Darstellung einer Route in Google Maps

Wenn es jedoch das Ziel ist, weitere Informationen wie z. B. in Abbildung 2 Points of Interests darzustellen, wird die Karte schnell unlesbar, da sich die Signaturen stellenweise überdecken, so dass diese weder eindeutig zu identifizieren sind noch zugeordnet werden kann, welche Points of Interests von einer Route aus gesehen bzw. besucht werden können.

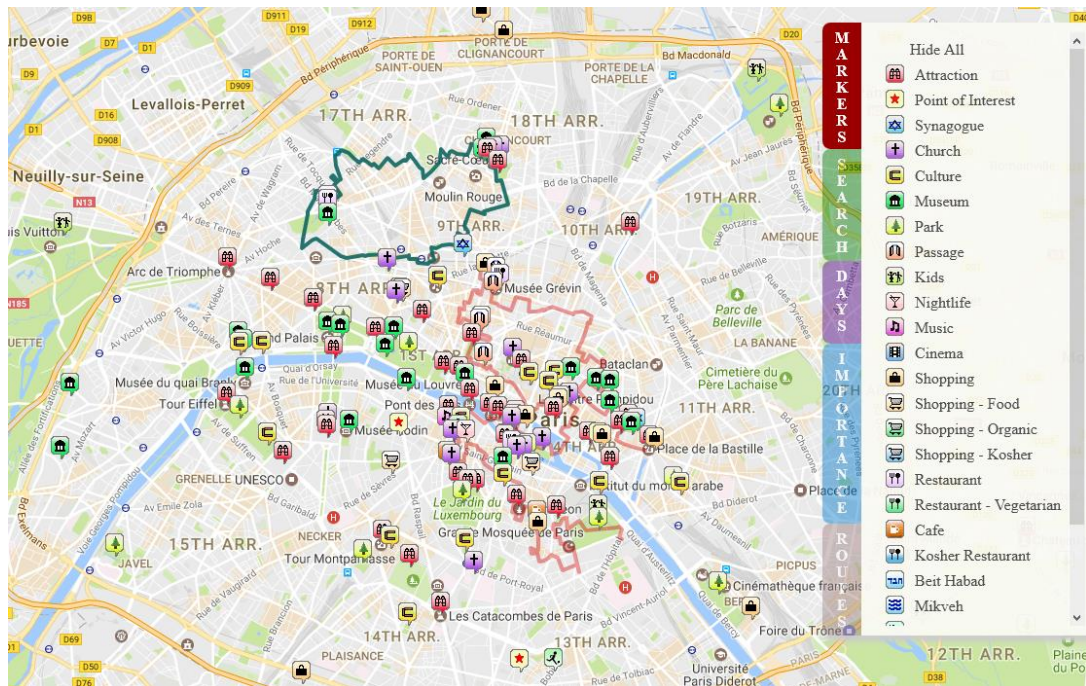


Abb. 2: Beispiel einer nutzergenerierten Karte mittels Google Maps

Dieses Problem besteht, da der Nutzer den darzustellenden Inhalt bestimmt, ohne dass überprüft wird, ob die daraus resultierende Karte den kartographischen Ansprüchen genügt. Des Weiteren hat der Nutzer keinerlei Möglichkeiten, die Standardkartenbilder zu bearbeiten, um beispielsweise für seine Thematik irrelevante Objekte auszublenden oder durch Verdrängung relevante Objekte freizustellen. Auch bei der Auswahl der Symbole gibt es meist keine Hilfestellung für die Entscheidung, ob eine Signatur geeignet ist um die gewünschten Informationen zu kommunizieren.

Im Rahmen dieser Arbeit wird ein konkreter Anwendungsfall, nämlich der Visualisierung von Meinungen und Gefühlen (Sentiments) über Orte betrachtet. Im Netz existieren zwar kaum Angaben mit expliziten Sentiments, jedoch sind auf sogenannten Travel Social Network Sites, oder auch Buchungsportalen für Urlaube eine Vielzahl von natürlichsprachlichen Bewertungen (Reviews) vorhanden, aus denen Sentiments extrahiert werden können. Diese Reviews dienen der subjektiven Beschreibung eines Ortes, ihre Darstellung erfolgt aber analog zu klassischen Reviews über Bücher oder Filme: Als Liste einzelner Reviews, die aus einem Titel gefolgt von dem Text sowie einer optionalen Bewertung (in Form von Sternen) bestehen und gelegentlich eine zusätzlichen Angabe enthalten, ob andere Nutzer dieses Review hilfreich fanden. Zwar sind durchaus Karten vorhanden, jedoch sind dort nur die Positionen der besprochenen Orte markiert und mit textuellen Beschreibungen verlinkt. Im Gegensatz dazu wird in dieser Arbeit die Integration der in den Texten enthaltenen Sentiments in die kartographische Visualisierung unter-

sucht. Hierzu wird einerseits analysiert, welche Charakteristiken die Daten aufweisen und welche unterschiedlichen Aspekte dem Kartenleser übermittelt werden können und welchen Einfluss dies auf die Wahl der Darstellungsart und der Signaturen hat. So dass aus diesen Erkenntnissen ein Softwaresystem entwickelt werden kann, was eine entsprechende Benutzerführung bzw. Unterstützung bei der Wahl der Darstellungsparameter erlaubt. Als zweiter Aspekt werden Verdrängungsalgorithmen für Punktsignaturen untersucht. Punktsignaturen sind eine Möglichkeit zur Darstellung von Sentiments. Da dort nicht nur die Existenz, sondern auch mindestens ein weiterer Wert für den Leser relevant ist, ist eine gute Lesbarkeit (keine Überlappung) zu gewährleisten.

1.1. Ziel dieser Arbeit

Im Rahmen dieser Arbeit werden mögliche kartographische Visualisierungen von ortsbezogenen Sentiments untersucht mit dem Ziel Erkenntnisse zu gewinnen, die für die Entwicklungen von Web Mapping Systemen hilfreich sind. Im Idealfall sollten diese Systeme jedem Benutzer ermöglichen, ansprechende und aussagekräftige kartographische Visualisierungen von beliebigen ortsbezogenen Informationen zu erzeugen. Im Folgenden werden zwei Aspekte für einen konkreten Anwendungsfall, der Visualisierung von Sentiments, untersucht:

- Möglichkeiten zur Unterstützung der Benutzer bei der Auswahl von geeigneten Symbolen.
- Bereitstellung von geeigneten Generalisierungsalgorithmen für die Visualisierung von Sentiments durch Punktsignaturen.

Für den ersten Aspekt werden zunächst verschiedene Darstellungsarten betrachtet und jeweils basierend auf der graphischen Semiologie von Bertin [BE11] Signaturen für Sentiments entworfen. Die Verwendung der graphischen Variablen bietet den Vorteil, dass die entsprechen Kriterien mit leicht verständlichen Fragen ermittelt und die Anpassung der Werte der Signaturen einfach berechnet werden kann. Ziel ist es herauszufinden, ob dieser Ansatz ausreichend ist um automatisch Signaturen zu generieren aus denen der Benutzer eine geeignete auswählen kann. Anschließend wird für eine Darstellungsart, der Visualisierung mittels eines zusätzlichen Layers mit Punktsignaturen, eine Evaluation mittels einer Umfrage in Bezug auf folgende drei Fragestellungen durchgeführt:

- Bei welcher Signatur wird die dargestellte Information korrekt ausgelesen?
- Sind die Benutzer in der Lage nur auf Grund der Symbole selbst ein für sie passendes Symbol auszuwählen?
- Ändern Benutzer ihre ursprüngliche Präferenz, nachdem sie mit den Symbolen gearbeitet haben? Wählen sie anschließend ein Symbol, dessen Werte sie häufiger richtig interpretiert haben?

Bei der Verwendung von Punktsignaturen für die Darstellung nutzergenerierter bzw. durch Nutzeranfragen ermittelter Daten ist das häufigste Problem, dass sich viele der Signaturen gegenseitig überlappen. Mögliche Generalisierungsoperationen zur Verbesserung der Lesbarkeit sind Selektion, Aggregation und Verdrängung. Des Weiteren kann durch Änderung der Signaturgrößen oder des Maßstabs eine Verbesserung der Darstellung erreicht werden. Da die meisten Web Mapping Systeme interaktiv gestaltet sind, ist letztgenannte Option bereits realisiert.

Der Schwerpunkt dieser Arbeit liegt auf der Entwicklung von Verfahren zur Verdrängung von Punktsignaturen. Ziel ist es dass die Verfahren zum Einem für Signaturen mit unterschiedlicher Form und Größe angewendet werden können und zum Anderen Anforderungen des Benutzers in Bezug auf die Gewichtung zwischen Lagegenauigkeit, Signaturgröße und Überlappungen durch für ihn nachvollziehbare Parameter berücksichtigt werden können.

I.2. Aufbau der Arbeit

In Kapitel 2 werden zunächst Voronoi-Diagramme vorgestellt, die als Hilfsstrukturen für die Verdrängungsverfahren verwendet werden. Kapitel 3 ist der graphischen Semiologie nach Bertin gewidmet. Anschließend folgt in Kapitel 4 eine kurze Zusammenfassung formaler psychologischer Modelle für Sentiments und eine Übersicht existierender Ansätze zur Visualisierung von Sentiments. Des Weiteren werden Verfahren zur Extraktion von Sentiments aus natürlichsprachlichen Texten vorgestellt. Einen Überblick über den Stand der Forschung in Bezug auf Generalisierungsverfahren für Punktsignaturen mit Schwerpunkt auf Verdrängungsverfahren folgt in Kapitel 5. Des Weiteren werden dort relevante kartographische Bedingungen und zugehörige Qualitätsmaße erläutert. In Abschnitt 5.3 wird die Lösbarkeit von Signaturkonstellationen untersucht und Kriterien zur Ermittlung von Überlappungen aufgestellt, die nicht durch Verdrängung aufgelöst werden können.

Kapitel 6 ist dem Entwurf der Signaturen gewidmet. Aufbauend auf Kapitel 3 werden Signaturen für drei verschiedenen Darstellungsarten erstellt: Anpassung gegebener Signaturen, Darstellung von Sentiments als Kontinua und Entwurf separater Punktsignaturen. In Kapitel 7 werden Verdrängungsverfahren für Punktsignaturen entworfen und implementiert, die auf Voronoi-Diagrammen bzw. auf Power-Diagrammen basieren. Ersteres ist für die Verdrängung gleichgroßer kreisförmiger Signaturen entwickelt worden und für die Verdrängung unterschiedlich großer Signaturen bzw. aus zwei Halbkreisen zusammengesetzter Signaturen wird jeweils eine Variante vorgestellt. Es werden zwei verschiedenen Heuristiken zur Bestimmung der Verschiebungsrichtung vorgeschlagen und untersucht. Des Weiteren wird eine Möglichkeit zur Steigerung der Effizienz durch Aufteilung der Punktmenge aufgezeigt. In Kapitel 8 werden zunächst die Ergebnisse der Umfrage zur Evaluation der Punktsignaturen vorgestellt. Anschließend wird das realisierte Verfahren für gleich große Kreissignaturen in drei Aspekten evaluiert: Grad der Reduzierung der Iterationsschritte durch Zerlegung der Punktmenge, erreichte Verminderung der Über-

lappungsfläche und Veränderung der relativen Lage der Punkte. Die beiden Varianten des Verfahrens werden bezüglich der Reduzierung der Überlappungsfläche mit den Ergebnissen verglichen die durch Verwendung des Voronoi-Diagramms bzw. von kreisförmigen Approximationen erzielt werden. Das 9. Kapitel schließt mit einer Zusammenfassung und einem Ausblick die Arbeit ab.

Voronoi-Diagramme

Ein Voronoi-Diagramm ist eine Partition des Raumes in Abhängigkeit einer gegebenen Punktmenge, die als Orte bezeichnet werden. Hierbei wird jeder Punkt des Raumes dem Ort zugeordnet, zu dem er den nächsten Abstand aufweist. Für jeden Ort s_i , ist seine Voronoizelle definiert als:

$$V(s_i) = \{x: |s_i - x| \leq |s_j - x| \text{ für alle } j \neq i\}.$$

Die Punkte die zu zwei Orten den gleichen Abstand aufweisen, bilden die Voronoikanten, d.h. eine Voronoikante wird durch zwei Orte definiert und liegt auf der Mittelsenkrechten der Verbindungslinie zwischen den beiden Orten. Punkte die von mindestens drei Orten gleich weit entfernt sind werden als Voronoiunkte bezeichnet. Hierbei handelt es sich um den Mittelpunkt des durch die drei Punkte definierten Kreises. Voronoi-Diagramme sind eine wesentliche Datenstruktur der Computational Geometry [Au91] die u. a. für Mustererkennung ([YHW14], [CS17]), Authentifizierung in Location Based Services [DSH+17] und Analyse und Visualisierung von Kriminalstatistiken verwendet werden [MFB17]. Darüber hinaus werden viele Varianten von Voronoi-Diagrammen für diverse Metriken ([DS15], [BLD+16], [BRW17]), für unterschiedliche Ortsgeometrien (Linien und Kreise [LD81], konvexe Polygone [Im96]) entwickelt und untersucht. Abbildung 3 zeigt ein einfaches zweidimensionales Voronoi-Diagramm für Punkte mit euklidischen Abstand (a), ein additiv (b), ein multiplikativ gewichtetes (c) und ein Voronoi-Diagramm für Linien (d). Bei additiven Voronoi-Diagramm werden von der Entfernung eines Punktes zu einem Ort positive Gewichte subtrahiert (In Abbildung 3(b) werden die Radien der Kreise als Gewichte verwendet) und bei multiplikativ gewichteten werden die Entfernungen zu einem Ort jeweils mit dem entsprechenden positiven Wert multipliziert (in Abbildung 3(c) sind die Gewichte als Zahlenwert gegeben).

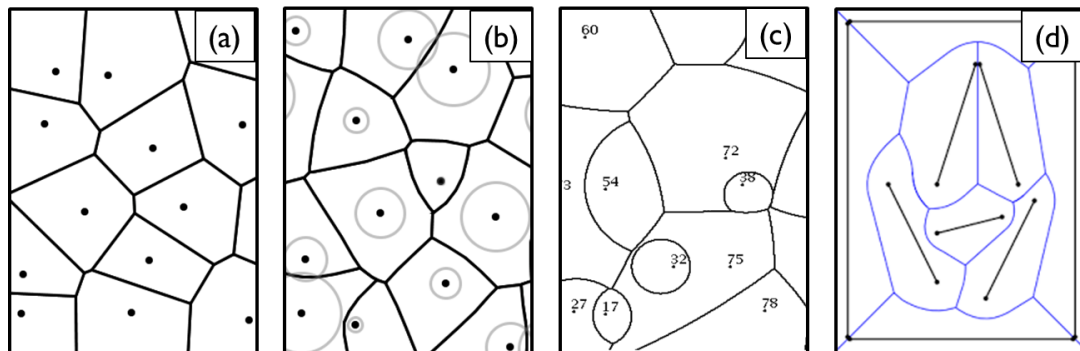


Abb. 3: Verschiedene Varianten von Voronoi-Diagrammen

Im Rahmen dieser Arbeit dient das Voronoidiagramm als Hilfstruktur für die Verdrängungsalgorithmen. Die relevanten mathematischen Eigenschaften werden in Abschnitt 2.1 aufgelistet. Neben dem einfachen zweidimensionalen Voronoi-Diagramm, werden auch Power-Diagramme (Abschnitt 2.3) als Hilfstruktur verwendet. Des Weiteren basiert eine Heuristik für die Verdrängungsverfahren auf Beobachtungen zum Schwerpunkt-Voronoi-Diagramm (Abschnitt 2.2).

2.1. Eigenschaften

Es folgt eine Auflistung bekannter mathematischen Eigenschaften (siehe z.B. [Ok92], [BKO+00]) die als Voraussetzung für die Verdrängungsverfahren benötigt werden. Im folgendem werden zwei Orte als direkte Nachbarn bezeichnet, wenn ihre Voronoizellen eine gemeinsame Kante aufweisen.

Bei Voronoizellen handelt es sich meist um konvexe Polygone. Die einzige Ausnahme, bilden die Zellen der Orte, die auf der konvexen Hülle der Orte liegen, da diese unbegrenzt sind. Bei der Verschneidung eines Voronoi-Diagramms mit einem konvexen Polygon sind alle Polygone der Ergebnismenge konvex. [SH76]

Auf Grund der Definition von Voronoi-Diagrammen entspricht der Abstand eines Ortes zu einer Kante, dem halben Abstand zwischen den beiden Orten. Dies bedeutet, dass die Berechnung der Abstände zwischen einem Ort und den Kanten seiner Voronoizelle hinreicht, um die Abstände zu seinen direkten Nachbarn zu berechnen.

Zur Bestimmung des nächsten Nachbarn eines Ortes müssen nur seine direkten Nachbarn betrachtet werden. Die Berechnung der beiden Orten mit dem geringsten Abstand aller nächsten Nachbarn, kann in linearer Zeit durchgeführt werden, da hierzu alle Kanten einmal betrachtet werden müssen und die Anzahl der Kanten sich linear zur Anzahl der Orte verhält.

Die Berechnung des Voronoi-Diagramms von n Orten kann mittels verschiedener Paradigmen, wie z.B. den Sweep-line-Paradigma [Fo87], oder mittels Divide and Conquer [SH75] in $O(n \log n)$ erfolgen.

Der duale Graph des Voronoi-Diagramms ist die Delaunay-Triangulation [De34]. Für die Delaunay-Triangulation gilt, dass im Umkreis eines Dreiecks kein Ort liegt (Umkreisbedingung).

2.2. Schwerpunkt-Voronoi-Diagramm (SVD)

Hierbei handelt es sich um spezielle Voronoi-Diagramme, die sich dadurch auszeichnen, dass die Orte der Voronoizellen gleichzeitig auch die Masseschwerpunkte ihrer Zelle sind. Zur Berechnung der Masseschwerpunkte ist eine Dichtefunktion notwendig. Im Rahmen dieser Arbeit werden nur SVD im \mathbb{R}^2 mit konstanter Dichtefunktion betrachtet. Eine gute Übersicht über SVD im allgemeinen bietet [DFG99].

Die bekannteste Methode, um aus einer beliebigen Punktmenge ein SVD zu erzeugen, ist Lloyds-Algorithmus [Ll82]. Hierbei handelt es sich um eine iterative Methode, bei der abwechselnd ein Voronoi-Diagramm der Punktmenge berechnet wird und anschließend die Orte auf die Schwerpunkte der Voronoizellen verschoben werden. Theoretisch endet die Iteration, wenn keine Änderung mehr notwendig ist, da die Schwerpunkte der Zellen mit ihren Orten übereinstimmen. Dass das Verfahren konvergiert und zu einem SVD führt, ist nur für den eindimensionalen Fall bewiesen [Fl64]. Für den zweidimensionalen Fall ist eine schwache globale Konvergenz nachgewiesen [EJR08]. In Abbildung 4 sind zwei verschiedene SVD mit jeweils 64 Orten abgebildet.

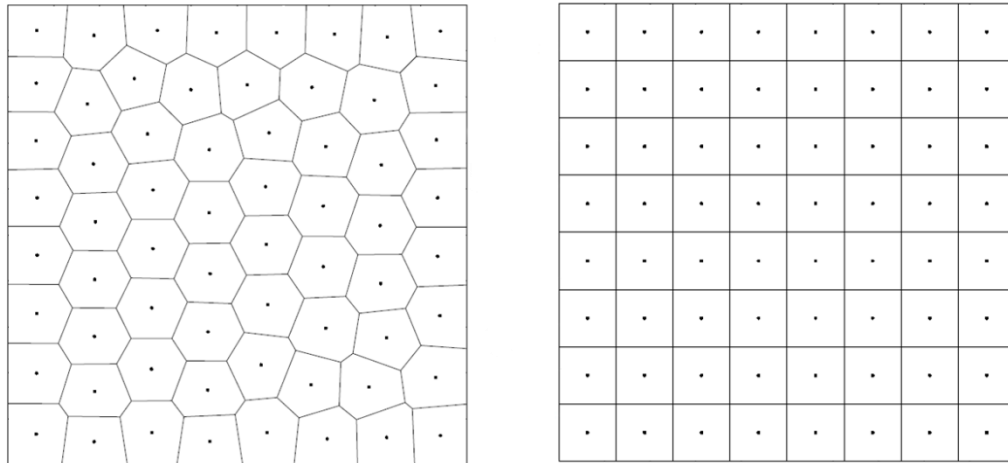


Abb. 4: Beispiele für Schwerpunkt-Voronoi-Diagramme

Bei praktischen Anwendungen wird die Iteration meist abgebrochen sobald der Abstand zwischen den aktuellen Orten der Voronoizellen und den Schwerpunkten ihrer Zelle eine festgelegte Distanz unterschreitet, so dass es sich bei dem Ergebnis strenggenommen nicht um ein SVD handelt. Allerdings wird in jedem Schritt die mittlere quadratische Distanz des Ortes der Voronoizelle zu allen anderen Punkten verkleinert und für die meisten Anwendungen ist diese Eigenschaft relevanter als die tatsächliche Übereinstimmung des Schwerpunktes mit den Orten. Die allgemeine Formel für die Berechnung dieser Distanz, die in der Literatur häufig auch als Energie bezeichnet wird, für eine Punktmenge $Z = \{z_i\}$ mit i von 1 bis n und eine Parkettierung $V = \{V_i\}$ mit i von 1 bis n lautet:

$$\varepsilon(Z, V) = \sum_{i=1}^n \int_{V_i} |y - z_i|^2 dy.$$

Bei der Punktverteilung mit minimaler Energie handelt es sich immer um ein SVD [DFG99] und die Orte sind gleichmäßig verteilt, d.h. die Abstände zwischen

benachbarten Orten sowie die Voronoizellen sind annähernd gleich groß [LYH+06].

2.3. Power-Diagramme

Eine Verallgemeinerung des Voronoi-Diagramms, bei dem es sich bei den Orten nicht um Punkte sondern um Sphären (im zweidimensionalen um Kreise) handelt, wird Power-Diagramm genannt [Au87]. Bei einem Power-Diagramm wird der Abstand zwischen einem Punkt x und einer Sphäre a mit Radius r , als quadratischer Abstand p des Punktes x zu dem Schnittpunkt t der Tangente von a durch x definiert (siehe Abbildung 5 I). Aufgrund des Satzes des Pythagoras lässt sich diese Distanz, der sogenannte Power-Wert p , durch $p = d^2 - r^2$ berechnen. Analog zum Voronoi-Diagramm wird jeder Punkt dann dem Kreis zugeordnet, für den der Power Wert p minimal ist. Schneiden sich zwei Kreise, so verläuft die Kante des Diagramms durch die beiden Schnittpunkte der Kreise (Abbildung 5 II). Liegt ein Kreis innerhalb eines anderen Kreises, so können beide Kreise auf einer Seite der durch sie definierten Kante liegen (Abbildung 5 III). Somit liegt, im Gegensatz zum Voronoi-Diagramm der Ort nicht immer innerhalb seiner zugehörigen Zelle und die Zelle eines Kreises kann auch leer sein.

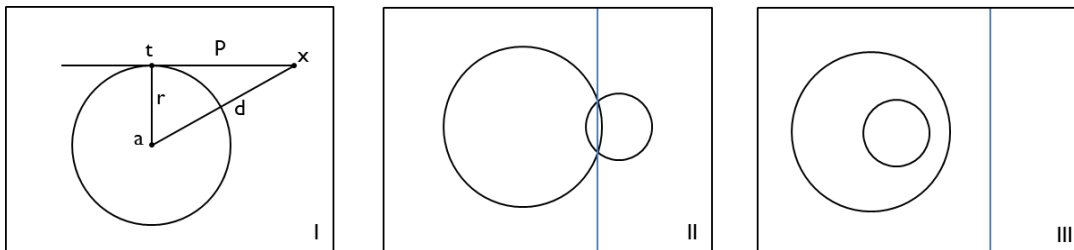


Abb. 5: Konstruktion von Kanten des Power-Diagramms

In Abbildung 6 ist ein Ausschnitt eines Power-Diagramms abgebildet bei dem die Zelle V_a den Kreis a nicht enthält und die Zelle V_b des Kreises B keine Fläche hat, da keine Punkte existieren für die der Power-Wert p zu B geringer ist als zu allen anderen Kreisen. Power-Diagramme können ebenso wie Voronoi-Diagramme in $O(n \log n)$ berechnet werden.

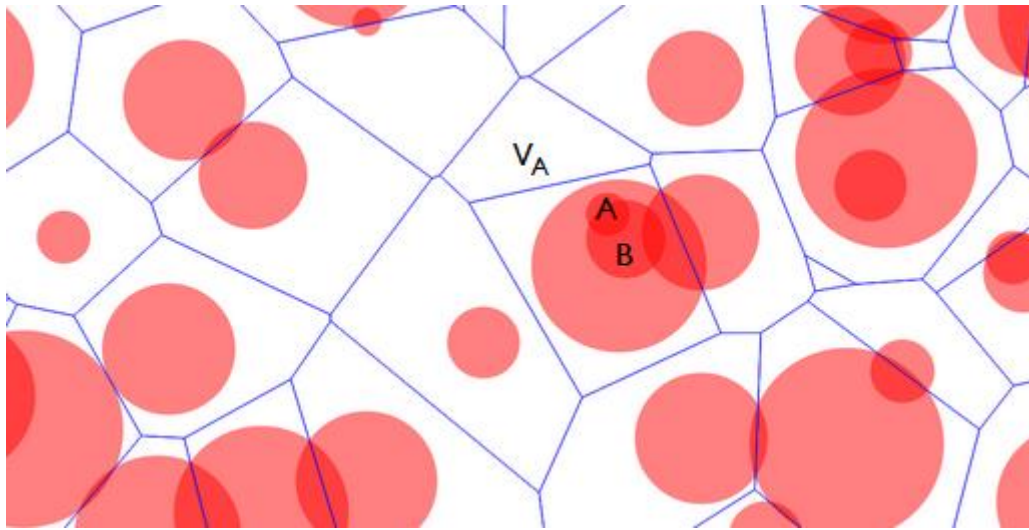


Abb. 6: Beispiel eines Power-Diagramms

Grundlagen der Informationsvisualisierung

Die kartographischen Visualisierungen, die im Rahmen dieser Arbeit erstellt werden, haben das Ziel Informationen über Sentiments effektiv zu kommunizieren. Bezogen auf die drei grundlegenden Paradigmen in der Kartographie: Karte als Bild (kommunikative/kognitive Tradition), Karte als Modell (analytische Tradition) und Karte als Absicht/Soziale Konstruktion [SH08], wird das zweite angewendet. Hierbei liegt der Schwerpunkt auf der Erstellung von automatisierbaren Regeln zum Entwurf bzw. Auswahl von Signaturen sowie der Bereitstellung von passenden Generalisierungsoperationen.

Bei Betrachtung der Karte als Bild sind die Grundelemente, die in Karten zur Darstellung von Geoobjekten verwendet werden, Punkte, Linien und Flächen, wobei die Wahl des Elements von der Objektgeometrie und dem Maßstab abhängt. Um zusätzliche Unterscheidung bezüglich der Objektarten und/oder Attributwerte von Objekten zu ermöglichen werden Signaturen verwendet. Signaturen können von bildhafter, symbolischer, geometrischer oder alphanumerischer Gestalt sein. Außerdem kann die dargestellte Information rein qualitativ sein, oder eine quantitative Komponente enthalten, die entweder stetig, gestuft oder mit Hilfe von Werteinheiten abgebildet wird. Für jede Gestalt sind in Abbildung 7 Beispielsignaturen für Punkte (lokal), Linien (linear) und Flächen dargestellt. [HGM02]

Als Grundlage für die in Kapitel 6 durchgeführten Entwürfe von Signaturen für Sentiments dient die graphische Semiologie von Bertin [Bel1], die in den folgenden Unterabschnitten vorgestellt wird. Neben den graphischen Variablen (Abschnitt 3.2) werden dort auch die Eigenschaften von Attributen (Skalen, Abschnitt 3.1) und Anwendungsszenarien betrachtet, die beeinflussen welche Variable bzw. Variablenkombination für die Darstellung eines Attributes bzw. von Attributkombinationen geeignet ist. Auf diese Weise können geeignete graphische Repräsentationen durch die Anwendung von grammatikalischen Regeln erstellt werden [BKCM+13]. Auf Basis dieser formalen Grundlage ist eine automatische Generierung von Visualisierungen, wie z.B. Diagrammen zur graphischen Darstellung von relationalen Daten möglich [Ma86]. In diesem Kontext ist ein Ranking der graphischen Variablen für die verschiedenen Skalen festgelegt. Jedoch existieren weitere unterschiedliche Rankings, die auf empirischen Studien basieren ([Ch84], [CM95]). Der Grund für die Unterschiede zwischen den Rankings ist, dass die Verwendbarkeit der graphischen Variablen, nicht nur von den Attributen abhängt, sondern auch von der Fragestellung die mit Hilfe der Visualisierung beantwortet werden soll ([NSH02]). Folglich ist eine Analyse der Aufgabenstellung, für die die Visualisierung entworfen wird, notwendig (Abschnitt 3.3).




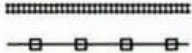
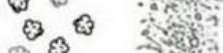

















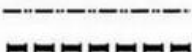
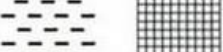











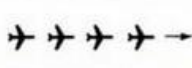

Form	Anordnung		
	lokal	linear	flächenhaft
Grundrissbild	  		
Bildhaft Aufrissbild	  		
Schrägbild	 		
Symbolisch	  		
Geometrisch	  		
Buchstabe, Ziffer, Zahl, Unterstreichung	  	<i>(unter dem Ortsnamen)</i> 	<i>sL 3 LÖ</i> <i>71 / 68</i>
Quantitäts- angabe	lokal und Signaturen- kartogramm	linear und Band- kartogramm	flächenhaft und Flächen- kartogramm
stetig mit Signaturenmaß- stab			
gestuft			
als Werteinheiten			

Abb. 7: Beispiele für Formen und Anordnungen von Signaturen[HGM02]

3.1. Charakterisierung der darzustellenden Information

In der graphischen Semiologie von Bertin wird eine Information in eine Invariante und eine oder mehrere Komponenten unterteilt. Die Invariante ist der Aspekt der gleich bleibt, während die Komponenten unterschiedliche Werte annehmen können. Welcher Teil einer Information als Invariante und welcher als Komponente eingestuft wird, ist eine Entwurfsentscheidung. So kann beispielsweise ein POI als Invariante und die Art des POIs, wie z.B. Theater, Restaurant, Denkmal, etc. als Komponente modelliert werden. Alternativ kann auch Theater als Invariante festgelegt werden und als Komponenten beispielsweise Anzahl der Sitzplätze oder Eintrittspreise. Für die Darstellung in Karten wird für die Invariante eine Signatur gewählt die in Abhängigkeit der Komponenten variiert wird. Für die Wahl einer passenden graphischen Variation, ist sowohl die Anzahl der verschiedenen Werte,

die eine Komponente annehmen kann relevant, als auch die Beziehung zwischen den Werten, den sogenannten Skalenniveaus:

- Nominalskala. Als Beziehung zwischen den Werten existiert nur Gleichheit oder Ungleichheit.
- Ordinalskala. Die Werte liegen geordnet vor, d.h. es ist möglich größer oder kleiner als Beziehung zwischen den Werten zu bestimmen.
- Kardinalskala, häufig weiter unterteilt in Intervallskala (Addition und Subtraktion möglich) und Verhältnisskala (Division und Multiplikation).

3.2. Graphische Variablen

Es gibt sechs graphische Variablen: Farbe, Form, Größe, Wert, Orientierung und Textur. Diese Variablen können einzeln oder kombiniert zur Darstellung von Komponenten einer Information verwendet werden. In Abbildung 8 sind jeweils für Punkte, Linien und Flächen verschiedene Variationen einer Variablen bei gleichbleibenden Werten der anderen fünf Variablen dargestellt.


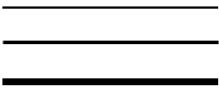
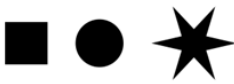
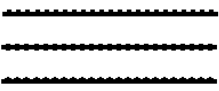




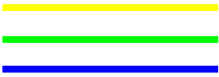
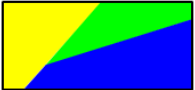

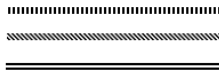




	lokal	linear	flächenhaft
Größe			
Form			
Wert			
Farbe			
Richtung			
Textur			

Abb. 8: Graphische Variablen

Bei Flächen können nur vier der sechs Variablen direkt angewendet werden, da die Größe und die Form der Fläche durch die Geometrie des Geoobjekts bereits festgelegt sind. Um diese beiden Variablen für ein flächenhaftes Objekt nutzen zu können, kann jedoch eine Punktsignatur verwendet werden, die entweder als Einzelsignatur oder als Muster in regelmäßigen Abständen innerhalb der Fläche platziert wird. Im folgendem wird für jede der Variable erläutert wie viele unterschiedliche Werte durch sie darstellbar sind, für welches Skalenniveau sie geeignet ist und welche weiteren Aspekte, wie z.B. Wechselwirkungen mit anderen Variablen und optische Täuschungen, zu beachten sind.

Größe: Bei einer Punktsignatur entspricht die Größe der Fläche der Signatur und bei einer Liniensignatur ihrer Breite. Mit Hilfe der Größe lassen sich nicht nur geordnete Komponenten darstellen, sondern in Abhängigkeit von der Form (bei Punktsignaturen) auch die Verhältnisse zwischen den Werten, wobei die Abschätzung der Größenverhältnisse mit zunehmender Komplexität der Signatur schwieriger wird, die Bestimmung der Ordnung der Signaturen jedoch nicht. Bei Liniensignaturen ist die Abschätzung der Größenverhältnisse einfacher, da nur die Ausdehnung in einer Dimension verändert wird. Aus diesem Grund wird bei länglichen Punktsignaturen (Balken) häufig zur Änderung der Größe nur die Höhe oder nur die Breite verändert, wodurch allerdings auch die Form der Signatur variiert. Die Anzahl der darstellbaren Werte hängt von dem Wertebereich ab. Falls die größte Signatur zehnmal so groß wie die kleinste ist, so lassen sich maximal 20 Stufen unterscheiden. Die Größe beeinflusst die Lesbarkeit der anderen Variablen, so dass bei der Wahl der minimalen Größe zu berücksichtigen ist, welche weiteren Variablen für die Darstellung verwendet werden.

Form: Es gibt eine unbeschränkte Anzahl verschiedener Formen, so dass Komponenten mit beliebiger Länge dargestellt werden können. Jedoch sind Formen im Allgemeinen nicht geordnet und somit nur für nominale Werte geeignet. Es lassen sich jedoch auch Gruppen von Formen definieren, für die sich eine Ordnung festlegen lässt, wie z.B. Häuser mit unterschiedlicher Anzahl von Stockwerken oder Fenstern oder Bäume mit unterschiedlicher Anzahl an Blättern, Blüten oder Früchten. Bei der Auswahl von Formen ist häufig zwischen einer einfachen Darstellung und einer einfachen Interpretation der Symbole abzuwägen. So können bildhafte und symbolische Signaturen intuitiv bestimmten Objektarten (z.T. auch konkreten Objekten) korrekt zugeordnet werden, benötigen jedoch meist mehr Platz, da ansonsten nicht alle Details lesbar sind. Geometrische Symbole sind meist auch in geringerer Größe lesbar, einfacher automatisch zu generieren, haben aber häufig keinen intuitiv eindeutigen Bezug zu den Objektarten, die sie darstellen.

Wert: Der Wert, auch Hellwert genannt, kann Werte zwischen 0 (schwarz) und 1 (weiß) annehmen. Bei schwarz-weiß Darstellungen (ohne Graustufen) entspricht der Hellwert dem Anteil an weißen Pixeln. Der Hellwert ist geordnet und kann auch für Kardinalsskalen verwendet werden. Allerdings ist es teils schwierig konkrete Werte abzulesen und die Wahrnehmung ist stark von der jeweiligen Umgebung abhängig. So wirken z.B. kleine Flächen dunkler, wenn sie von größeren weißen Flächen umgeben sind. Ein Beispiel für eine weitere optische Täuschung ist in Abbildung 9 zu sehen. Die Balken mit den gleichen Ziffern weisen jeweils den gleichen Hellwert auf, werden aber auf Grund des Hintergrundes als unterschiedlich wahrgenommen. Weiterhin ist zu berücksichtigen, dass sich der Hellwert bei Änderung der Farbe, Textur und Richtung ebenfalls ändern kann. Der Hellwert einer Farbe ist bei Verwendung des HSV-Farbraumes als V-Wert angegeben. Bei feinen Texturen und Strukturen zur Richtungsangabe wird ein einheitlicher Hellwert für die Gesamtfläche wahrgenommen, der dem gewichteten Mittelwert der vorkommenden Hellwerte entspricht. Bei groben Texturen ist eine visuelle Bestimmung eines Hellwerts der Gesamtfläche meist nicht möglich.

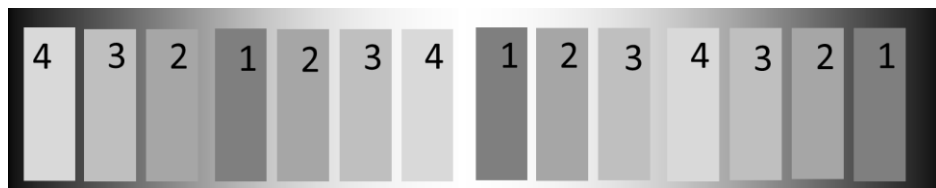


Abb. 9: Einfluss des Hintergrundes auf Wahrnehmung des Hellwerts

Farbe: Die Farbwahrnehmung bei Menschen ist individuell verschieden und zusätzlich von der Umgebung, der Beleuchtung und dem Material bzw. dem Medium abhängig. Der Mensch kann ca. 10 Millionen verschiedene Farben wahrnehmen, und in der Computergraphik wird die Anzahl meist auf 16,8 Millionen begrenzt (2^{24}). Zur Repräsentation von Farben gibt es verschiedene Farbräume. Die gebräuchlichsten sind:

- RGB: Additive Farbmodell mit den Grundfarben Rot, Grün und Blau, das vor allem bei digitalen Displays verwendet wird.
- CMYK: Subtraktives Farbmodell mit den Grundfarben Cyan, Magenta, Gelb und Schwarz, das beim Drucken verwendet wird.
- HSV: Die Farben werden durch die Parameter Farbwert, Sättigung und Hellwert definiert. Die Parameter entsprechen mehr der menschlichen Farbwahrnehmung in dem Sinne, dass der Zusammenhang zwischen Parameter- und Farbänderung intuitiv verständlich ist

Farben sind in erster Linie für nominale Attribute geeignet, wie z.B. grau für Wohngebäude und rot für öffentliche Gebäude. Allerdings rufen einige

Farben starke Assoziationen zu geordneten Eigenschaften hervor, wie beispielsweise Rot - Orange/Gelb - Grün zu negativ – neutral – positiv (Am-pelschema). Des Weiteren werden Rot- und Brauntöne meist als warme Farben und Blautöne als kalt wahrgenommen, so dass sich diese zur Darstellung von geordneten Werten wie z.B. der Temperatur eignen. Auch ist die Verwendung von Farbverläufen für geordnete Werte üblich, jedoch wird in diesen Fällen nicht nur die Farbe sondern auch der Hellwert variiert und dieser bestimmt die wahrgenommene Ordnung. Hierbei ist zu beachten, dass der Vergleich der Hellwerte schwieriger wird je größer der wahrgenommene Unterschied zwischen den Farben ist. Ferner gibt es nicht nur die Assoziationen zwischen Attributen und Farben, sondern auch zwischen Objektarten und Farben. So werden, z.B. Gewässer mit Blautönen, Wälder und Wiesen mit Grüntönen assoziiert, so dass eine entsprechende Farbgebung die Interpretation von Darstellungen erleichtert. Allerdings ist dabei zu beachten, dass die Zuordnung von Farben zu Eigenschaften und Objekten kulturell geprägt ist [NE04].

Richtung: Bei Punktsignaturen ist der Wertebereich der unterscheidbaren Richtungen von der Form abhängig. Bei unsymmetrischen Formen sind Richtungen zwischen 0° und 360° möglich. Kleine Winkeländerungen lassen sich einfacher unterscheiden je größer und länglicher die Form ist. Bei Linien und Flächen wird die Richtung über eine Schraffur realisiert, so dass eine Verwendung der Richtung immer auch Einfluss auf den Hellwert und die Textur der Signatur hat. Die Richtung kann für alle Skalenniveaus verwendet werden.

Textur: Die Anzahl an verschiedenen Texturen ist unbegrenzt, jedoch ist die konkrete Anzahl, die für eine Signatur verwendet werden kann, von ihrer Größe abhängig. Bei Texturen wird zwischen groben/rauen und feinen/glatte Texturen unterschieden, so dass eine Ordnung definiert werden kann. Allerdings ist die Anordnung von Texturen, die stark unterschiedliche Strukturen aufweisen, verhältnismäßig schwierig. Für Linien- und Flächensignaturen ist die Verwendung von Textur und Richtung zur Darstellung unterschiedlicher Komponenten nicht empfehlenswert, da Schraffuren, die sich nur durch die Richtung der Linien unterscheiden, dennoch als verschiedene Texturen wahrgenommen werden. Des Weiteren ist zu beachten, dass bei Texturen ein Vibrationseffekt auftreten kann, d.h. es entsteht der Eindruck, dass eine Fläche flimmert. Dies geschieht insbesondere bei schwarz-weiß Signaturen mit einem Hellwert um 0,5 und Strukturen, die größer als 1mm sind.

3.3. Wahrnehmungsarten von graphischen Variablen

Es gibt eine Reihe verschiedener Fragestellungen, die mit Hilfe einer kartographischen Darstellung beantwortet werden können, z.B.:

- Was befindet sich an einem bestimmten Ort?
- Wo befinden sich bestimmte Objekte?
- Wie ist die Anordnung aller Objekte mit einer bestimmten Eigenschaft?
- Wie ist die Anordnung aller Objekte einer Objektart mit unterschiedlichen Eigenschaften?
- Wie ist die Rangfolge von Objekten bezüglich einer Eigenschaft?

Wie gut die Verwendung einer graphischen Variablen zur Darstellung für die Beantwortung einer Fragestellung ist, hängt von den Wahrnehmungsarten ab, die die Variable unterstützt. Es wird zwischen assoziativer, dissoziativer und selektiver Wahrnehmung unterschieden.

Assoziative Wahrnehmung: Wenn Signaturen bezüglich einer assoziativen Variable unterschiedliche Werte aufweisen, werden sie dennoch als zu einer Gruppe gehörig gesehen, solange die Werte der anderen Variablen übereinstimmen. Form, Richtung, Farbe und Textur sind assoziative Variablen.

Dissoziative Wahrnehmung: Wenn sich Signaturen bezüglich einer dissoziativen Variablen unterscheiden, werden sie als unterschiedlich wahrgenommen. Unterschiedliche Werte von dissoziativen Variablen führen zu einer Änderung der Sichtbarkeit der Signaturen. Hellwert und Größe sind dissoziativ.

Selektive Wahrnehmung: Alle Signaturen mit dem gleichen Wert einer graphischen Variable können schnell identifiziert werden. Ob eine Variable eine selektive Wahrnehmung unterstützt, hängt zum einen von der Anzahl der unterschiedlichen Werte und zum anderen von der Unterscheidbarkeit der verwendeten Werte ab. Farbe ist am besten für die selektive Wahrnehmung geeignet, die Anzahl der verwendbaren Farben hängt von dem Hellwert der Farben und von der Größe der Objekte ab. Soll eine selektive Wahrnehmung möglich sein, so ist die Anzahl der Werte beim Hellwert auf sieben und bei Textur und Größe auf fünf zu beschränken. Die Orientierung ist nur bei Punktsignaturen bei bis zu vier unterschiedlichen Werten, und bei Liniensignaturen bei maximal zwei unterschiedlichen Werten selektiv. Die Form unterstützt keine selektive Wahrnehmung.

Weitere Eigenschaften, die zur feineren Differenzierung der Anforderungen an die Darstellung herangezogen werden können, sind, ob nur die Eigenschaften von benachbarten Objekten zu vergleichen sind, ob der absolute oder der relative Wert einer visualisierten Eigenschaft erkannt werden soll und wie strikt die Anforderungen bezüglich der Genauigkeit und der Vollständigkeit der Darstellung ist.

Sentiments

Die in dieser Arbeit für die kartographische Visualisierung gewählte Thematik ist „Sentiment“, da Sentiments zum einem im Bereich der Analyse von nutzergenerierten Inhalten ein relevantes Forschungsgebiet sind (Sentiment Analysis, Opinion Mining) und zum anderen in Anwendungsgebieten wie Tourismus, Gastronomie und Stadtplanung der Ortsbezug der Sentiments relevant ist. Es gibt eine Reihe von Projekten (z.B. Bio Mapping [No09], WiMo [MWK09], EmoMap [OH11] und EmBaGIS [BZP+11]), die sich mit ortsbezogenen Sentiments, bzw. Emotionen/Gefühlen oder Polarität beschäftigen. Im Rahmen dieser Arbeit wird Sentiment als Oberbegriff der drei anderen Begriffe aufgefasst. Im folgendem werden zunächst einige bekannte formale psychologische Modelle für Sentiments vorgestellt. Anschließend werden grundsätzliche Ansätze zur Sentimentextraktion beschrieben und ein Überblick gegeben, welche Verfahren für räumliche Daten angewendet werden. Es folgen zwei Abschnitte in denen verschiedene Visualisierungen für Sentiments, nämlich allgemeine (Abschnitt 4.3) und kartographische (Abschnitt 4.4), präsentiert und diskutiert werden.

4.1. Psychologische Modelle

Es existieren zwei grundlegende Ansätze zur Modellierung von Gefühlen: Dimensionsmodelle und diskrete Modelle. Bei den diskreten Modellen besteht Konsens über sechs Basisemotionen: Freude, Überraschung, Trauer, Wut, Angst und Ekel. Diesen Gefühlen entspricht in jedem sozialen und kognitiven System eine eindeutig unterscheidbar Repräsentation unabhängig von Ethnie und Zivilisationsstufe [EF71]. Bei zusätzlicher Berücksichtigung von neuropsychologischen Effekten ergeben sich vier weitere eindeutig zu identifizierende Gefühlszustände: Interesse, Verachtung, Scham und Schuld [Iz77]. Nach einer psychoevolutionären Theorie werden neben den sechs Basisemotionen auch Vertrauen und Wachsamkeit als grundlegende Emotionen festgelegt. Des Weiteren ergeben sich dort durch Betrachtung der Intensität der Emotionen, zu jeder dieser acht Basisemotionen zwei zusätzliche, wie z.B. Zorn als intensive Wut und Verärgerung als abgeschwächte Form [PI01].

Das Ziel der dimensional Modelle ist es, eine Anzahl (k) von Basis-Emotionspaaren zu identifizieren, wie z.B. angenehm-unangenehm oder aktiv-gedämpft, die einen k -dimensionalen Raum definieren. In frühen Modellen wurden die folgenden drei Dimensionen identifiziert (z.B. [OMM75]):

- Wertigkeit: Wie angenehm ist das Gefühl?
- Stärke: Wie gut lässt sich das Gefühl kontrollieren?

- Aktivierung: Wie erregend ist das Gefühl?

Jedoch gibt es in der Psychologie keinen Konsens über die Anzahl sowie die Natur der Dimensionen und neben den dreidimensionalen Modellen werden sowohl zweidimensionale Modelle (Wertigkeit und Aktivierung [Ru80]), als auch höherdimensionale Modelle, wie z.B. ein vierdimensionales Modell, in dem die Unvorhersehbarkeit als vierte Dimension hinzugefügt wird [FSR+07], diskutiert.

Der Vorteil von dimensional Modellen ist die einfache Berechnung von Gefühlen sofern die Werte der Grunddimensionen bekannt sind. Des Weiteren ist eine Aggregatfunktion von mehreren Gefühlen einfach zu bestimmen, da für die einzelnen Dimensionen jeweils die Mittelwerte gebildet werden können und sich somit eine neue Emotion ergibt. Allerdings ist die Zuordnung von Emotionsbezeichnungen zu Gebieten nicht eindeutig festgelegt. Ein weiterer Kritikpunkt bei klassischen dimensional Modellen ist die fehlende Möglichkeit sich widersprechenden Gefühle zu modellieren [BTK15]. Grund hierfür ist die Annahme, dass das Affektsystem (Gefühlsempfinden) bipolar ist. Im Gegensatz dazu ist eine Kombination von Basisemotionen zur Darstellung von gemischten Gefühlen möglich und es lassen sich eindeutig verschiedene Gesichtsausdrücke für einige Kombinationen identifizieren [DTM14].

4.2. Extraktion aus Texten

Verfahren zur Sentimentanalyse können in folgende Klassen unterteilt werden, wobei die Verwendung der Begriffe in der Literatur nicht einheitlich ist:

- Polaritätsbestimmung: Aussagen werden als positiv, neutral oder negativ klassifiziert, wobei auch die Stärke angegeben werden kann. Für positiv werden meist Werte zwischen 0 und 1 und für negativ Werte zwischen 0 und -1 verwendet.
- Bestimmung von Sentiments (dimensionales Modell): Aussagen werden in Bezug auf die jeweiligen Dimensionen, meist Wertigkeit und Aktivierung, bewertet.
- Bestimmung von Sentiments (diskretes Modell): Jeder Aussage werden eine oder mehrerer Emotionen zugeordnet.
- Stance Analyse: Es wird untersucht ob eine Aussage Ablehnung oder Zustimmung zu einer anderen Aussage signalisiert.

Für jede dieser Analyseaufgaben können zwei konzeptionell unterschiedliche Ansätze verfolgt werden: maschinelle Lernverfahren oder Wörterbuch-basierte Verfahren.

Die gängigsten maschinellen Lernverfahren (ML) für die Sentiment-Analyse sind Naives Bayes, Support Vector Machines (SVM) und Maximum Entropie (ME) [PLV02]. Welcher Ansatz bzw. welches Verfahren für eine Sentiment-Analyse ge-

eignet ist, hängt unter anderem von der Domäne ab. Für räumlichen Daten, werden sowohl maschinelle Lernverfahren als auch wörterbuchbasierte angewendet, die jeweils ähnliche Werte bzgl. Precision und Recall erreichen. In Abschnitt 4.2.1 wird eine Übersicht der Sentiment-Analyse-Verfahren gegeben, die von aktuellen peer-reviewed Arbeiten verwendet werden, die sich mit der Darstellung von Sentiments beschäftigen. In Abschnitt 4.2.2 folgt eine Betrachtung von wörterbuchbasierten Methode, da sie für die Berechnung der Sentiments im Rahmen dieser Arbeit verwendet werden.

4.2.1. Sentiment-Analyse von räumlichen Daten

Bei den hier vorgestellten Arbeiten handelt es sich um eine Auswahl von Beiträgen, die sich auch mit der Darstellung von Sentiments befassen und die über den SentimentVis Browser [KPK17] abfragbar sind. Grund hierfür ist, dass im Rahmen dieser Arbeit kein eigenes Extraktionsverfahren entworfen wird, sondern aus den bereits verfügbaren Verfahren (ein) geeignete(s) in Abhängigkeit von der Visualisierung gewählt werden soll.

In Tabelle 1 ist angegeben wie oft welcher Ansatz für welche Analyse-Aufgabe verwendet wird. Da die Stance-Analyse in nur in einer Arbeit betrachtet wird [MSK+17], ist sie nicht in der Tabelle aufgeführt. Die meisten Arbeiten beschränken sich auf die Polaritätsbestimmung (23 von 28), in drei der Arbeiten ([WWL+10], [DPI6],[PKS16]) werden zusätzlich noch Sentiments mittels des dimensional Modells bestimmt. Für die Polaritätsbestimmung werden häufiger Wörterbuch-basierte Methoden angewendet. Die Auswahl der Wörterbücher erfolgt in erster Linie an Hand der Sprache und der Art der Texte. Da oft keine speziell für die Thematik der räumlichen Daten erstellten Wörterbücher existieren, werden sie teils manuell um die häufigsten Wörter [DPI6], oder um die Bewertung von Emoticons ergänzt [YWC+16], oder auf einen Teil der Wörter, z.B. auf Adjektive, beschränkt [WWL+10]. In Bezug auf ML-Verfahren werden sowohl SVM [LWW+16], als auch Naive Bayes [MBB+11] und ME [SBB+15] verwendet. Teils werden auch Wörterbücher mit ML kombiniert ([PKS16], [CGG+15], [LHW+15]). Bei der Verwendung diskreter Sentimentmodelle werden entweder die sechs Basisemotionen ([ZKK+09] WB, [CWV+17] ML), oder eine Teilmenge betrachtet ([ZDW+12] NB, [CWV+17] ML).

Tabelle 1: Häufigkeit von Sentiment-Analyse-Verfahren

	Wörterbuch (WB)	ML	WB+ML
Polarität	12	8	3
Dimensionales Sentiment-Modell	2	1	1
Diskretes Sentiment-Modell	1	2	1

4.2.2. Wörterbuch-basierte Verfahren.

Der zentrale Teil der Wörterbuch-basierten Ansätze ist eine Abbildung von Stimmungswerten auf einzelne Wörter oder Kombinationen von Wörtern. Abhängig von der Ebene, auf der die Analyse durchgeführt werden soll, werden die Werte für alle Wörter innerhalb eines Dokuments oder innerhalb eines Satzes oder für alle Wörter, die einen Aspekt beschreiben, aggregiert. In den meisten Fällen ist die Aggregationsfunktion entweder die Summe oder der Mittelwert. Eines der am häufigsten verwendeten Wörterbücher ist das Ergebnis der Affective Norms for English Words (ANEW) Studie [BL99], die Valenz-, Erregungs- und Dominanzwerte für über 1.000 englische Wörter enthält. Ein anderes Modell wird von SENTIWORDNET verwendet, ein Wörterbuch, das für jede WORDNET Synset drei numerische Werte, einer basierend auf, wie negativ die enthaltenen Wörter sind, einer auf, wie positiv sie sind und der dritte auf, wie neutral sie sind. Eine detaillierte Beschreibung dieses Ansatzes findet sich in [ES06]. Ein anderes Wörterbuch ist AFINN [Ni11], das ähnlich wie ANEW ist, aber das speziell für den Einsatz auf Tweets entwickelt wurde. Alle drei sind zumindest für Forschungszwecke frei verfügbar. Die meisten Wörterbücher wurden manuell erstellt, aber es gibt auch einige halbautomatische Ansätze, die auf einer kleinen manuell annotierten Wortmenge basieren, die entweder mit Suchmaschinen [TL03] oder mit Synonymen und Antonymen erweitert wurden ([KMM+04], [Li12]).

Neben dem Wörterbuch mit den Sentiment Werten sind weitere Wörterbücher für Negationen und Betonungsworte relevant, um zufriedenstellende Ergebnisse zu erzielen [BCM+12]. SentiStrength ist ein Beispiel für eine Methode, die diese drei Arten von Wörterbüchern verwendet und zusätzlich eine Liste mit Abkürzungen und Slangwörtern [TBP+10] enthält. Die Sentiment Werte reichen von -5 bis +5, wobei 0 nicht verwendet wird. Die Ausgabe entspricht dem positivsten und negativsten Stimmungswert innerhalb des analysierten Textes. Ein sehr ähnliches System ist der Lexicon-Based Classifier [PT12]. Allerdings liefert er als Ergebnis nur eine Klasse, die entweder positiv, negativ oder neutral ist. Eine weitere Wörterbuch-basierte Methode ist der Semantic Orientation CALCulator (SOCAL) [TBT+11]. Im Gegensatz zu den anderen Methoden werden Worte unterschiedlich gewichtet in Abhängigkeit von ihrer Wortart. Daher müssen die Texte mit Hilfe eines Part of Speech Taggers vorverarbeitet werden, z.B. Constraint Grammatik [KVH+95], Viterbi Algorithmus [Fo73] oder Brill Tagger [Br92]. Der Sentiment Wert jedes Wortes wird unter Berücksichtigung der zugehörigen Betonungswörter und Negative berechnet. Danach werden die Werte aggregiert.

4.3. Visualisierung von Sentiments

Für die Visualisierung von Sentiments gibt es eine Vielzahl verschiedener Ansätze, wie z.B. die Repräsentation in Linienplots, Pixel-Matrizen, Node-Link-Strukturen, Clouds/Galaxien, Text, Icons oder Karten. Welcher dieser Ansätze möglich bzw. sinnvoll ist hängt sowohl von den Daten, der durchgeführten Sentimentanalyse

(siehe Abschnitt 4.2) als auch der Aufgabe/Fragestellung ab, die mit Hilfe der Visualisierung bearbeitet werden soll. In Bezug auf die Daten sind ihre Domäne, die Quelle sowie grundlegende Eigenschaften, wie z.B. ob es sich um Daten mit Zeit- oder Raumbezug oder Daten zur Beschreibung von Netzwerken handelt relevant. Mögliche Aufgabenstellungen sind die Detektion auffälliger Regionen, die Klassifikation bzw. Clustering von Sentimentwerten, der Vergleich, Übersicht oder Überwachung. Des Weiteren kann auch die Darstellbarkeit von Unsicherheiten oder bei interaktiven Visualisierungen die Möglichkeiten zur Navigation relevant sein. Einen umfassenden Überblick über den Stand der Forschung in Bezug auf all diese Optionen bietet [KPK17]. Es folgt ein kurzer Überblick über die dortigen Beobachtungen in Bezug auf die Visualisierungen von Sentiments mit Raumbezug bezüglich folgender Fragestellungen:

- Für welche Aufgaben werden Kartendarstellungen gewählt?
- Welche Graphischen Variablen werden für welches Sentimentmodell verwendet?
- Welche graphischen Variablen werden in den Kartendarstellungen verwendet?

Häufig dient die Karte zur Übersicht über die Verteilung der Daten [PKS16], zur Navigation [LWW+16] oder für beides [MSK+17], [CWV+17] ohne dass Informationen zu den Sentimentwerten auf der Karte visualisiert werden. In einigen Arbeiten wird auch vollständig auf eine Kartendarstellung verzichtet ([WWL10], [ZDW+12], [WFL+13], [WGS15], [FZC+17]). Bei der Verwendung von Graphen [ZKK+09] oder Diagrammen [YWC+16], sind diese meist auf der Karte verlinkt und werden in einem separaten Fenster angezeigt.

Da es sich bei den Systemen meist um interaktive Visualisierungen handelt, können positive und negative Sentiments getrennt visualisiert werden ([LHW+15], [QRQ15], [SBB+15], [DPI6]). Die Darstellung erfolgt meist über einfache Signaturen, wie Marker ([MBB+11], [SQC16]), oder Punkte ([CSS+14], [CGG+15], [HHM+16]). Als zusätzliche Information wird häufig noch die Anzahl der Dokumente bzw. Tweets über die Größe der Signatur visualisiert ([SBB+15], [SHW+12], [WFL+12]). Wird zusätzlich noch die Thematik oder der Nutzertyp visualisiert, werden unterschiedliche Formen für die Signaturen verwendet ([WWL+10], [CLS+12]).

Die Sentiments werden bis auf eine Ausnahme, die die Textur zur Darstellung verwendet [KTL+14], durch Farben dargestellt. Für die Darstellung der Polarität ist die Verwendung des Ampelschemas (Rot für Negativ, Gelb/Orange für Neutral und Grün für Positiv) weit verbreitet ([CLS+12], [SHW+12], [CDG+14], [CGG+15], [WGS15]), wobei für Neutral auch Grau ([HRJ+13], [HHM+16]), Weiß [SQC16] oder Beige [LWW+16] verwendet wird. Alternativ werden auch warme Farben (Rot, Orange, Gelb) für positive Werte und kalte Farben (Blau) für negative Werte verwendet ([WFL+12], [PKS16], [YWC+16]), aber auch gegen-

sätzlich Farbuordnungen (Blau für Positiv und Rot für Negativ) sind nicht selten ([MBB+11], [CSS+14], [LHW+15], [SBB+15], [LSB+16]).

Für die Visualisierung diskreter Sentimentwerte bieten sich auch komplexe bildhafte Signaturen an, wobei diese meist in psychologischen Arbeiten zur Erfassung von Sentiments und nicht für ihre Visualisierung verwendet werden. Am Häufigsten ist dort die Verwendung von Gesichtern. In Abbildung 10 ist links eine Comickarstellung der sechs Basisemotionen, sowie auf der rechten Seite zwei Beispiele für eine manuelle Kombination der Gesichtszüge zur Darstellung gemischter Emotionen. Die Berechnung der Gesichtstrukturen kann auch automatisch aus diskreten Sentiments [DTM14], oder Mittels Wertigkeits- und Aktivierungswerten erfolgen [GAP+10].

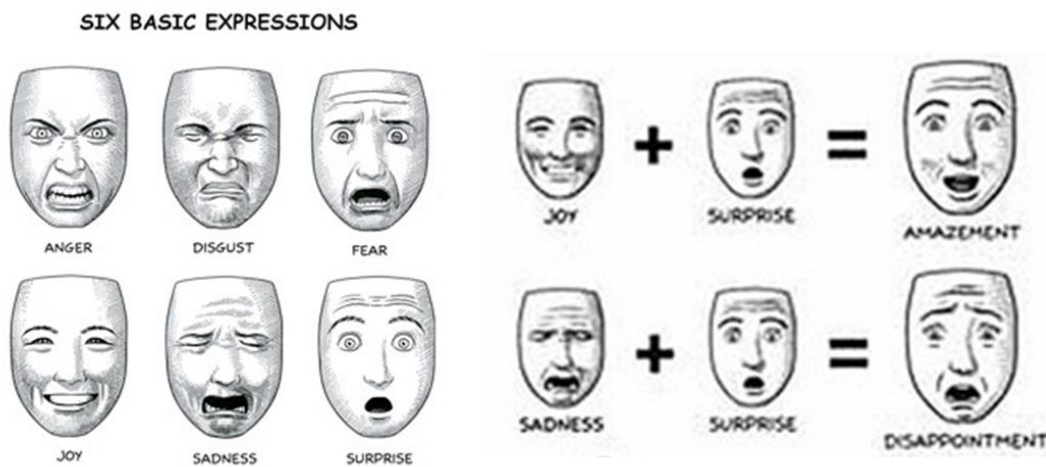


Abb. 10 Gesichtsausdrücke zur Visualisierung von Basisemotionen und gemischten Emotionen [Mc06]

Eine weitere komplexe aber durchaus intuitiv nachvollziehbare Visualisierung ist in Abbildung 11 dargestellt. Diese Visualisierung ist im Rahmen einer Studie entstanden bei der die Probanden jeweils angeben, an welchem Stellen im Körper sie Gefühle wahrnehmen.

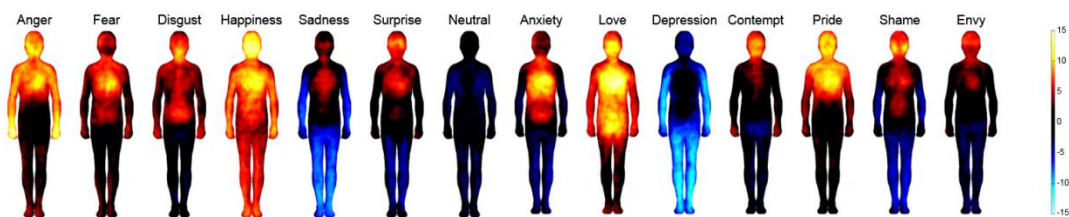


Abb. 11 Bodily Maps von Emotionen [NGH+13]

4.4. Kartographische Darstellung

Die Einstellung von Menschen zu ihrer Umgebung hängt nicht nur von objektiv messbaren Parametern wie z.B. dem Wetter ab, sondern auch von subjektiven Empfindungen (Sentiments). Die Bedeutung dieser subjektiven Umgebungsqualität ist nicht nur für ihr Wohlbefinden ([UI84], [Ka93]), sondern auch für die Entscheidungen bei der Auswahl von Wegen [Za01] und sogar für soziale Interaktion [SKD04] durch mehrere Forschungsarbeiten belegt. Die Ausbreitung von Smartphones und die wachsende Akzeptanz von Web 2.0 Technologien bietet nunmehr eine Möglichkeit, um die Wahrnehmung bzw. Bewertungen der Benutzer über ihre Umgebung zu sammeln. Diese Daten können einen zusätzlichen Nutzen für eine Reihe von Applikationen in verschiedenen Bereichen darstellen wie z.B. für die Stadtplanung [AW11], den Tourismus oder die Fußgängernavigation [KG12]. Der Fokus dieser Forschungsarbeiten liegt dabei zumeist auf der Erfassung gefühlsbezogener Informationen und darauf, wie diese Informationen in bestehende Stadtplanungs- und Navigationssysteme integriert werden können und nicht auf der kartographischen Darstellung.

Zur Visualisierung von Sentiments, die die Umgebung betreffen, gibt es nur relativ wenige Arbeiten, die sich explizit mit dem Entwurfsprozess beschäftigen, jedoch einige die gelungene Visualisierungen präsentieren: Für eine der ersten digitalen Karten zur Visualisierung von Gefühlen wird zunächst die gefühlte Sicherheit von Einzelpersonen erfasst, indem sie die Karte eines Stadtgebiets unterschiedlich einfärben. Für die Wahl der Farben wird das Ampelschema verwendet, d.h. grün kennzeichnet die Gebiete, in denen sich die befragten Personen sicher fühlen, rot Angst einflößende Gebiete, orange alle anderen bekannten Gebiete und blau unbekannte Gebiete. Ein Beispiel einer solchen Comfort Map für Los Angeles ist in Abbildung 13 auf der linken Seite zu sehen. Anschließend werden alle Comfort Maps in einer Karte zusammengefasst, in der Gebiete in denen das Gefühl der Sicherheit höher als der Mittelwert ist grün eingefärbt sind und Gebiete in denen es niedriger ist in rot, wie in Abbildung 12 auf der rechten Seite dargestellt. [MBQ01]

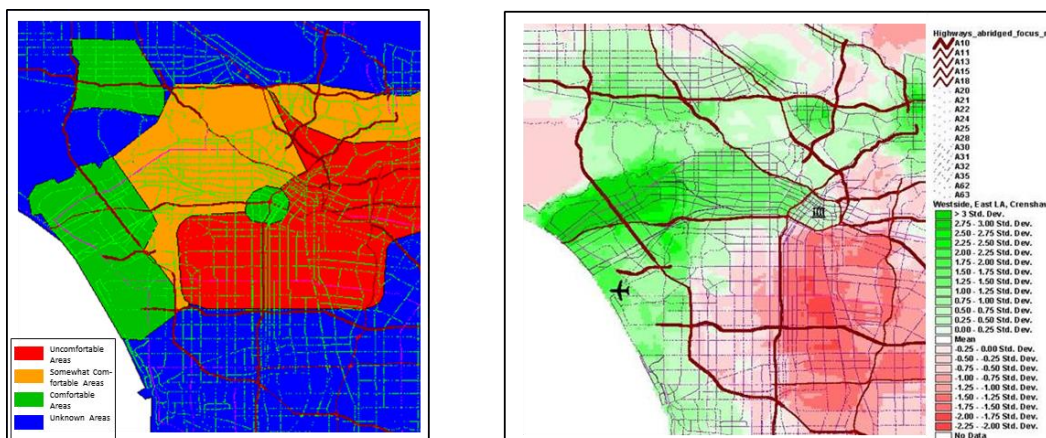


Abb. 12 Comfort Maps: Einzeldarstellung und aggregierte Darstellung [MBQ10]

Eine Reihe weiterer Visualisierungen sind im Rahmen des Biomapping-Projekts entstanden [No09]. Diese Karten wurden von Teilnehmern im Rahmen verschiedener Workshops u. a. in Paris, San Francisco und Stockport (Stadt in England) erstellt. Die gefühlsbezogene Information wurde nicht durch explizite Eingabe der Teilnehmer gesammelt, sondern durch Messung der Hautspannung (Galvanic Skin Response) während die Teilnehmer durch eine Stadt gegangen sind. Anschließend wurden die gesammelten Daten diskutiert und textuelle Beschreibungen wurden zu den erstellten Karten hinzugefügt, woraus mehrere unterschiedliche Visualisierungen resultierten. Aus drei der auf diese Weise entstandenen Karten ist jeweils ein Ausschnitt in Abbildung 13 dargestellt. Zum Beispiel wird in der Karte für Stockport die emotionale Anspannung (der gemessene Wert) durch Säulendiagramme in verschiedenen Farben visualisiert. Das Farbspektrum reicht von Schwarz zu rot und auf der Hintergrundkarte sind Skizzen von allem, was den Teilnehmern wichtig war, festgehalten (siehe Abbildung 13, links). Für die Karte von Ost-Paris wurde eine schematische Darstellung der abgetretenen Wege gewählt, auf der Anspannung durch rote Punkte symbolisiert ist (siehe Abbildung 13, mitte), und die Karte von San Francisco besteht ausschließlich aus roten bis schwarzen Punkten, die den Gemütszustand der Teilnehmer repräsentieren (siehe Abbildung 13, rechts). Die Visualisierungen sind das Ergebnis einer kreativen Gruppenarbeit, wobei die Motivation und die Gründe für die Wahl der Darstellung in der Arbeit nicht weiter diskutiert werden. Es wird zwar eine große Bandbreite an möglichen Visualisierungen aufgezeigt, die jedoch manuell und intuitiv entstanden und nicht automatisiert reproduzierbar sind.

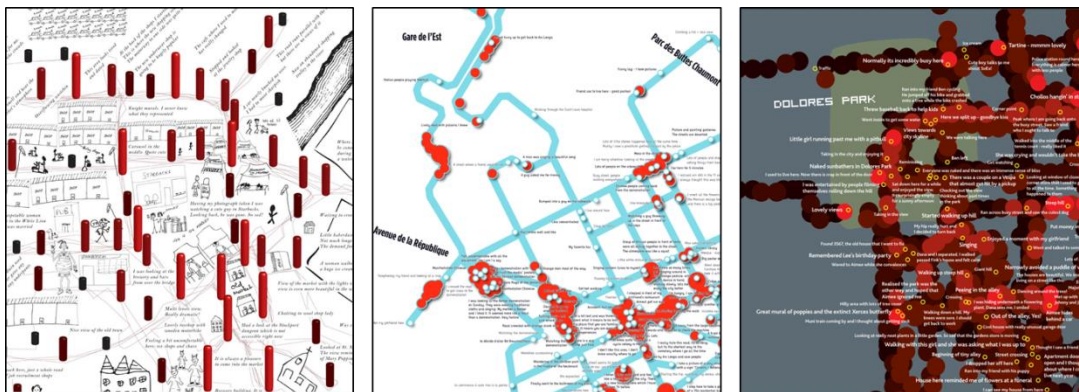


Abb. 13 Ausschnitte aus Karten des Biomapping-Projekts [No09]

Es besteht auch die Möglichkeit vorhandene Web 2.0 Plattformen zu nutzen um Sentiments darzustellen. So können mit Hilfe des Yahoo! Placemaker Web Services Reiseberichte kartographisch dargestellt werden [AL10], oder auch mit Google Maps die in Nachrichtenartikeln enthaltenen Sentiments als Säulendiagramme visualisiert werden [KFK+08]. Ein weiteres Beispiel für die Verwendung mit Google Maps in der die Sentiments mit Hilfe von Smilies dargestellt werden ist Tourpedia [CDG+14]. Zur Auflösung von Clutter in höheren Zoomstufen werden anstelle

der Signaturen Zahlen dargestellt die angeben wie viele Signaturen sich in dem Bereich befinden, wobei der Bereich nicht dargestellt wird, wie in Abbildung 14 links zu sehen. Allerdings sind die Überlappungen je nach gewählten Bereich und Zoomstufe recht hoch (siehe Abbildung 14 Mitte). In diesen Arbeiten steht jedoch nicht die Visualisierung an sich im Vordergrund, sondern ihre technische Umsetzung.

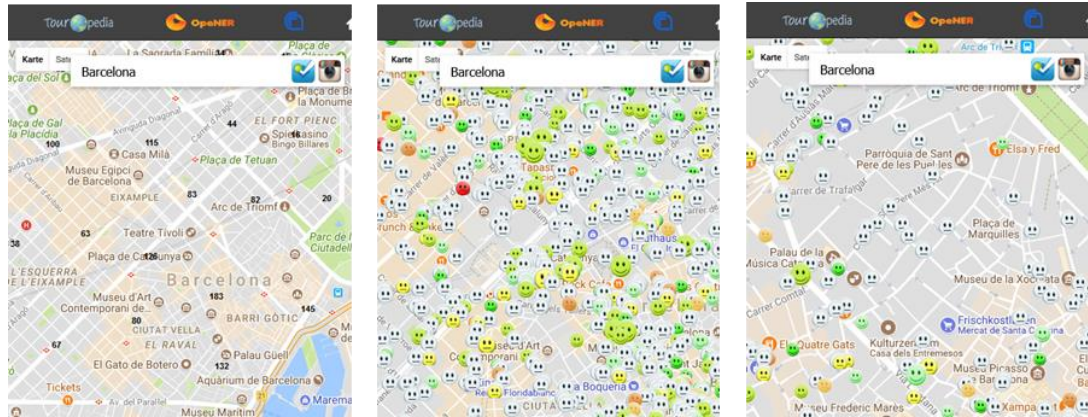


Abb. 14 Visualisierung von Sentiments durch Smilies in verschiedenen Zoomstufen [CDG+14]

Im Gegensatz dazu liegt der Fokus dieser Arbeit auf dem systematischen Entwurf von Signaturen für Sentiments basierend auf objektiven kartographischen Kriterien und der Möglichkeit diese, unabhängig von einem konkreten System, automatisiert zu visualisieren.

Generalisierung von POI-Karten

Die Verwendung komplexer Punktsignaturen in kartographischer Visualisierungen von benutzergenerierten Inhalten bietet dann einen Mehrwert für den Benutzer, wenn die Darstellung lesbar und insbesondere die Sichtbarkeit der Signaturen gewährleistet ist. Da in vielen Anwendungsfällen die minimale Größe der Signatur größer als die symbolisierte Fläche ist und POIs häufig geclustert auftreten, sind Überlappungen zwischen Signaturen wahrscheinlich. Zur Auflösung solcher Darstellungskonflikte können verschiedene Strategien aus der kartographischen Generalisierung eingesetzt werden [BW10].

- a) Inhaltserhaltend: Ziel ist es, möglichst alle Punkte zu erhalten und die Lesbarkeit durch Verdrängung sicher zu stellen. Es wird angenommen, dass die Daten schon gefiltert vorliegen und die Darstellung der Objekte wichtiger ist als ihre exakte Position.
- b) Interaktiv: Die Daten werden aggregiert dargestellt und dem Benutzer wird ermöglicht mit der Karte zu interagieren, so dass die Informationen über die einzelnen Komponenten der Aggregationen verfügbar bleiben.
- c) Positionserhaltend: Die Lage der dargestellten Objekte ist wichtiger als der Erhalt aller Objekte. Als Generalisierungsoperationen werden hauptsächlich Selektion (absolute Lage) und Typifizierung (relative Lage) verwendet.

In dieser Arbeit werden Informationen zu Orten aus Texten im Hinblick auf gegebene Fragestellungen extrahiert. Somit sind die zu visualisierenden Informationen bereits nach Relevanz gefiltert. Die Visualisierungen sind für touristische Fragestellungen ausgelegt, jedoch nicht für Navigationszwecke oder Notfallsituationen. Somit ist die exakte Lage der POIs weniger relevant als ihre thematischen Attribute. Des Weiteren sollen thematische Informationen graphisch dargestellt werden und eine sinnvolle Aggregation komplexer Signaturen führt meist zu einem Informationsverlust, oder einer Erhöhung der Komplexität [Mel6]. Folglich ist Strategie a) im Kontext dieser Arbeit am angemessensten und dieses Kapitel widmet sich der Betrachtung von Verdrängungsverfahren.

Trotz des Bedarfs existiert keine etablierte Methode zur Verdrängung von Punktdaten. Ein Grund hierfür ist das Punktdaten bisher eine geringe Rolle in der kartographischen Generalisierung gespielt haben. In traditionellen Karten sind Punktsignaturen entweder das Ergebnis einer vorangegangenen Generalisierung (durch Symbolisierung) oder sie werden für Objekte verwendet, die nur in großmaßstäbigen Karten (z.B. Liegenschaftskarten) angezeigt werden. Infolgedessen entstehen Konflikte für gewöhnlich zwischen Punktdaten und linienförmigen oder flächenhaften Objekten und nicht zwischen verschiedenen Punktsymbolen.

5.1. Verdrängung

Ziel der Kartengeneralisierung ist die Vereinfachung der Darstellung geographischer Daten, um eine gut lesbare Karte in einem festgelegten Maßstab zu erzeugen [Ru08]. Meist wird die Generalisierung im Kontext von Maßstabsänderungen, von größeren auf kleinere Maßstäbe, betrachtet. Dabei werden zwei widersprüchliche Ziele verfolgt; zum einem die Datenmenge zu reduzieren, zum anderen die Ähnlichkeit zur ursprünglichen Karte beizubehalten [PAH+09].

Die Herausforderung bei der Generalisierung von POI-Karten liegt bei der Reduzierung der Überlappungen der POI-Signaturen (Cluttering). Die Hintergrundkarte ist bei POI-Karten vorgegeben und wird während der Generalisierung nicht verändert. Für Cluttering existieren verschiedene Definitionen, in denen Aspekte wie Dichte und Layout der POIs, Anwendung, Hintergrundkarten und Leser unterschiedlich stark berücksichtigt werden [MS14]. Als Generalisierungsoperationen werden in existierenden Systemen in erster Linie Selektion und Aggregation verwendet [HG12]. Die Selektion kann entweder auf Grundlage der thematischen Attribute der POIs erfolgen oder auf Grund der räumlichen Anordnung der POIs. Bei der thematischen Selektion wird unterschieden, ob diese an Hand inhaltsbasierter [LGS10], kollaborativer [DK10], demographischer [RRS10], wissensbasierter [FFJ+10], community-basierter [RRS11] oder hybrider Kriterien [Bu02] durchgeführt wird. Zusätzlich können bei der Selektion noch die benachbarten (konkurrierenden) POIs sowie der Ort des Benutzers berücksichtigt werden [MSV15]. Für die geometrische Selektion sind eine Reihe von iterativen Ansätzen und Clustering-basierten Ansätzen entwickelt worden [BBC+04]. Bei iterativen wird eine Teilmenge der POIs ausgewählt und anschließend durch Heuristiken iterativ die Ähnlichkeit zu der ursprünglichen räumlichen Anordnung der POIs verbessert. Bei der Clustering-basierten Selektion werden zunächst Cluster gebildet und jedes Cluster durch einen repräsentativen Punkt ersetzt. Des Weiteren können auch Voronoi-Diagramme für die Auswahl der POIs, die nicht dargestellt werden, verwendet werden [YW08].

Für die on-the-fly-Generalisierung von POI-Karten für mobile Geräte kann bei Verwendung von hierarchischen Datenstrukturen jeder Baumebene ein Level of Detail (LOD) und dies dann einem oder mehreren Maßstäben zugeordnet werden. Auf diese Weise wird der Zugriff auf die für den jeweiligen Maßstab relevanten Daten beschleunigt und ein schnelles Zoomen ermöglicht [BPE04]. Neben Selektion, Vereinfachung und Aggregation kann auch die Verdrängung auf Basis eines Quadtree definiert werden [BW13a]. Hierbei ist für jede Quadtree-Ebene festgelegt, wieviele POIs ein Knoten enthalten darf. Falls diese Anzahl überschritten wird, werden die POIs auf einen geometrisch benachbarten Knoten verschoben, sofern er noch nicht die maximale Anzahl an POIs enthält. Bei GIS-Systemen, in denen mehrere Generalisierungsoperationen ausgeführt werden (z.B. [EBW05], [BW13a]), wird die Verdrängung üblicherweise zum Schluss ausgeführt, da sie die Anzahl der dargestellten Objekte nicht weiter verringert [BW10].

Für die Bewertung von Generalisierungsoperationen für konkrete Anwendungsfälle von POI-Karten existiert eine Reihe von Kriterien die teils aus der Kartographie und teils aus der Informationsvisualisierung stammen [KA13]. In Bezug auf die primären Kriterien für Sentiment Maps (siehe Kap. 6.1): Vermeidung versteckter Symbole, Erhalt von Attributwerten, Zugriff auf individuelle Symbole und Lokalisierbarkeit ist die Verdrängung am besten geeignet, sofern die Gesamtanzahl der POIs im gewählten Ausschnitt und Maßstab darstellbar sind. Neben den speziell für das Auflösen von Cluttering entworfenen Verfahren, existiert eine Vielzahl von Methoden, die ebenfalls für die Verdrängung von POIs verwendet werden können. Diese lassen sich in inkrementelle Verbesserungen und ganzheitliche Ansätze einordnen. In einigen inkrementellen Methoden wird eine Menge vorher festgelegter Positionen getestet und z.B. lokale Suche ([MP01], [FS04]) oder Simulated Annealing [WJ98] verwendet, um eine zufällig oder heuristisch generierte Anfangskonfiguration zu verbessern. Außerdem wurden einige deterministische Verfahren entwickelt, wie z.B. ein Greedy-Ansatz, bei dem in jedem Schritt die beiden Objekte verdrängt werden, die an der größten Überlappung beteiligt sind [Ru98], oder die Anwendung eines Gradientenverfahrens [LJ01].

Bei ganzheitlichen Ansätzen werden mehrere Kartenobjekte gleichzeitig betrachtet. Hierfür können beispielsweise Cluster-Analysen verwendet werden, um alle Objekte die von einem Konflikt betroffen sind zu identifizieren. Anschließend werden sie von Konfliktmittelpunkt radial nach außen verdrängt. Auf diese Weise werden die Abstände zwischen den sich überlappenden Objekten vergrößert ohne die relative Lage der Objekte zueinander zu verändern [Ma94]. Allerdings werden die Objekte weiter verschoben als es zur Auflösung der Überlappung notwendig ist. In anderen Studien wird das Verdrängungsproblem als ein Gleichungssystem formuliert, dass durch eine Optimierungsmethode gelöst werden kann, wie z.B. die kleinste Quadrate Methode ([Ha99], [SK99], [Se05]), Finite Elemente Methode [Hø00] oder unter Verwendung sogenannter Snakes [BM97]. In [AO02] wird für die Gebäudegeneralisierung mit Hilfe des Delaunay-Skeletts eine Aufteilung des Raumes berechnet, die dem Voronoi-Diagramm ähnelt. An Hand dieser Aufteilung wird für jedes Gebäude ein Verschiebungsvektor berechnet, wobei die Verschiebungsvektoren auch auf die benachbarten Gebäude wirken. Handelt es sich bei den Objekten ausschließlich um Punkte, so kann direkt das Voronoi-Diagramm für die Berechnung der Verschiebungsvektoren verwendet werden.

Da bei der Verdrängung von Signaturen für Sentimentkarten, keine Real-Time-Anforderungen bestehen und es sich ausschließlich um Punktdaten handelt, bietet sich ein Voronoi-Diagramm-basierter Ansatz an. Ein Vorteil bei der Verwendung von Voronoi-Diagrammen liegt darin, dass der Raum den Objekten zugeordnet wird und somit nicht nur die Konflikte, sondern auch die zur Verfügung stehende Fläche bei der Berechnung der Verschiebungsvektoren berücksichtigt werden kann. Des Weiteren ist keine Beschränkung auf mögliche Verschiebungspositionen notwendig, wie bei inkrementellen Ansätzen. Aus diesem Grund wird im Folgen-

den ein iteratives Verdrängungsverfahren entwickelt, bei dem der Verschiebungsvektor jedes POIs auf Basis seiner Voronoizelle berechnet wird.

5.2. Kartographische Bedingungen und Qualitätsmaße

Um die Qualität von kartographischen Darstellungen zu bewerten, sind zunächst die Anforderungen an diese Darstellungen zu spezifizieren und anschließend geeignete Qualitätsmaße zu definieren [Jo98]. Da die Anforderungen sowohl vom Anwendungsfall, als auch vom Inhalt und Maßstab der Karte sowie dem Trägermedium abhängig sind, existieren in diesem Bereich zahlreiche Betrachtungsweisen mit unterschiedlichen Schwerpunkten. So werden in [GM95] die Anforderungen an räumliche Daten allgemein betrachtet und Genauigkeit (bzgl. Position, Attribute und Semantik der räumlichen Objekte), Vollständigkeit, logische Konsistenz, Lineage und zeitliche Information als wichtige Elemente zur Beurteilung der Qualität von Geodaten identifiziert. Ein anderer Ansatz ist es, die Anforderungen an die Daten und ihre Darstellung explizit im Generalisierungsprozess zu integrieren [HW07]. Hierzu werden Constraints formuliert, die sich auf die Position, Topologie, Form, Struktur, Funktion und Lesbarkeit der Karte beziehen ([RP96], [WD98], [Ha03]). Dagegen werden in [SW07] die Anforderungen in Abhängigkeit der Beziehungen der beteiligten Objekte klassifiziert:

- Horizontal: Beziehungen zwischen verschiedenen Objekten innerhalb eines Maßstabs, sowie Eigenschaften eines Objekts.
- Vertikal: Beziehungen zwischen dem gleichen Objekt in verschiedenen Maßstäben, bzw. Änderung der Eigenschaften eines Objektes
- Update: Änderung eines Objektes in der Realität und daraus resultierende Änderung der Darstellung.

Für das in dieser Arbeit entwickelte Verdrängungsverfahren sind lediglich horizontale Beziehungen relevant, da es zur Erstellung einer Graphik in einem festgelegten Maßstab verwendet wird. Da es sich bei den darzustellenden Informationen um Punkte bzw. Punktmengen handelt, ist neben der Lagegenauigkeit und Lesbarkeit insbesondere die räumliche Verteilung relevant [BWI1].

Die Erhaltung der Lagegenauigkeit steht im Widerspruch zur Auflösung der durch die Symbolisierung auftretenden Überlappungen der Punktsignaturen. Um eine gewisse Lagegenauigkeit zu garantieren, wird eine maximal zulässige Abweichung definiert und als Constraint schon während der Verdrängung berücksichtigt.

Als Maß für die Lesbarkeit der einzelnen Punktsymbole bietet sich der Flächeninhalt der Überlappungen von Signaturen an, die durch die Symbolisierung auftreten. Allerdings lässt ein Vergleich der Gesamtüberlappungsfläche nur bedingt Rückschlüsse auf die Verbesserung der Lesbarkeit der Karte zu. Dies liegt daran, dass Überdeckungen von weniger als einem Viertel der Signaturfläche bei (symmetrischen) Punktsymbolen keinen negativen Einfluss auf die Genauigkeit oder die Geschwindigkeit bei der Interpretation der Symbole haben [KHA14]. Somit kann eine

Darstellung mit größerer Überlappungsfläche lesbarer sein, wenn für mehr Signaturen die Überdeckungsfläche jeweils geringer ist. Aus diesem Grund wird als Qualitätsmaß in Bezug auf die Lesbarkeit der Karte die Anzahl der Signaturen, bei denen die Überlappungsfläche größer als 25% bzw. größer als 50% ist, verwendet. Für die globale Lesbarkeit einer Karte sind die Menge der dargestellten Information und ihre räumliche Verteilung relevant [HSD15]. Bei POI-Karten kann die Menge der Information durch die Anzahl der Punkte pro Fläche bestimmt werden, bzw. unter Berücksichtigung der räumlichen Verteilung als Anzahl der Punkte pro Cluster und der für das Cluster zur Verfügung stehenden Fläche.

Durch die Verdrängung von Punkten ändert sich nicht nur die Lage der Punkte sondern gegebenenfalls auch die relative Lage der Punkte zueinander bzw. zu anderen Kartenobjekten und folglich auch die räumliche Anordnung der Punkte. Um diese Veränderung zu bestimmen, können Methoden der Punktmuster-Analyse verwendet werden [OU03]. Wie gut die zugrundeliegende räumliche Punkteverteilung bei der Generalisierung erhalten bleibt, kann beispielsweise durch den visuellen Vergleich von Kerndichteschätzungskarten (Kernel Density Estimation, KDE) bestimmt werden [BW13b]. Hierzu wird die KDE-Funktion auf die darzustellenden Punkte angewendet, um eine geglättete Oberfläche zu schätzen, die die Dichte der Punktverteilung darstellt. Die Funktion lautet:

$$f(x, y) = \frac{1}{nh^2} \sum_{i=1}^n k\left(\frac{d_i}{h}\right)$$

Hierbei ist $f(x,y)$ der Dichtewert an der Stelle (x,y) , n die Anzahl der Punkte, h die Bandbreite (auch Suchradius genannt), d_i die euklidische Distanz zwischen dem Punkt i und der Stelle (x,y) und k ist eine Dichtefunktion, die auch Kern (kernel) genannt wird. Als Parameter sind der Kern und die Bandbreite zu wählen, wobei als Kern häufig ein Polynom vierten Grades gewählt wird ([BG95], [Ra02], [Le04]). Für die Wahl einer geeigneten Bandbreite gibt es verschiedene Ansätze, deren Eignung von den Anwendungsgebieten und den Daten abhängig ist [Tu93]. So wird etwa in ArcGis die Bandbreite h standardmäßig auf:

$$h = 0.9 \min \left(SD, \sqrt{\frac{1}{\ln 2}} D_m \right) n^{-0.2}$$

gesetzt, wobei SD die Standardentfernung, D_m der Median und n die Anzahl der Punkte ist.

Des Weiteren existieren eine Reihe von Maßen, die für kartographische Anwendungen entwickelt wurden und sich auf Graphen (Delaunay-Triangulationen) [WF94] oder Voronoi-Diagramme als Hilfsstrukturen stützen ([LH02], [XTS08], [HS10]). Zur Bewertung eines Verdrängungsverfahrens für Punkte bietet sich ein

Entropie-basiertes Maß an, das ursprünglich für den Informationsgehalt von Karten entwickelt wurde [LH02]. Durch dieses Maß lassen sich unterschiedliche räumliche Anordnungen unterscheiden, jedoch werden geringfügige Änderungen der Position der Punkte nicht detektiert [HS10]. Zur Berechnung wird das Voronoi-Diagramm der Punkte konstruiert und die Entropie der Fläche der Voronoizellen bestimmt durch:

$$HI = \frac{\sum_{i=1}^n p_{vi} \log p_{vi}}{\log \frac{1}{n}}$$

wobei p_{vi} die relative Größe der Voronoizelle v_i ist und n die Anzahl der Punkte.

Zur Evaluation des Verdrängungsverfahrens werden in Abschnitt 8.2 drei horizontale Maße verwendet: der Flächeninhalt der Überlappungen zwischen Signaturen zur Bestimmung der Lesbarkeit, KDE zum visuellen Vergleich und die Entropie der Voronoizellenfläche als quantitatives Maß für die räumliche Anordnung der Punkte.

5.3. Lösbarkeit von Überlappungen

Neben Maßen zur Beurteilung der Qualität einer generalisierten Abbildung sind im Kontext von nutzergenerierten Karten auch Kriterien relevant, an Hand derer eine Aussage getroffen werden kann, ob eine gegebene Punktmenge im gewählten Maßstab überhaupt sinnvoll darstellbar ist. Ein mögliches Kriterium hierfür ist die Anzahl der nicht lösbaren Überlappungen. Denn auf Grund der Beschränkung der Distanz, um die ein Symbol verschoben werden darf, sind nicht alle Überlappungen auflösbar. Die Bestimmung der exakten Anzahl der nicht lösbaren Überlappungen ist NP-hart [MS91]. Dies liegt daran, dass unter Umständen durch die Verdrängung einer Signatur zwar eine Überlappung aufgelöst werden kann, aber dadurch eine neue Überlappung entsteht. Jedoch ist es bei der Verwendung gleichgroßer kreisförmiger Signaturen möglich, durch die lokale Betrachtung einer Überlappung festzustellen, ob sie nicht auflösbar ist und somit eine untere Schranke für die Anzahl der nicht auflösbaren Überlappungen zu berechnen. Hierzu folgt zunächst eine Betrachtung verschiedener Zählweisen für die Anzahl der Überlappungen, bevor eine effiziente Methode zur Bestimmung der unteren Schranke vorgestellt wird.

Die Anzahl der Überlappungen bezeichnet, wörtlich genommen, die Anzahl der Schnittflächen, wobei die Unterteilung in einzelne Schnittflächen entweder rein geometrisch (disjunkte Flächen), oder unter Beachtung der Thematik (am Schnitt beteiligte Objekte) getroffen werden kann. Diese Interpretation ist insbesondere sinnvoll, wenn die Schnittflächen relevant für die Anwendung sind, d.h. wenn es sich um real existierende Überlagerungen verschiedener Thematiken handelt, wie z.B. Bodenarten, Bodenfeuchte und Landnutzung. Hier sind die Überlappungen jedoch nur durch die Darstellung verursacht und beeinträchtigen die Lesbarkeit der sich überlappenden Symbole. Folglich ist die Anzahl der Symbole, die an Über-

lappungen beteiligt sind, aussagekräftiger als die Anzahl der Schnittflächen. Betrachtet man nicht nur die Gesamtanzahl der Objekte, die sich überlappen, sondern für jedes Objekt die Anzahl der Objekte, mit denen es überlappt, ist eine feinere Differenzierung möglich, die jedoch für die Bewertung der Darstellbarkeit nicht relevant ist, wie im Folgendem an Hand eines Beispiels erläutert wird. Unter der Bedingung, dass jede Signatur an mindestens einer Überlappung beteiligt ist, lassen sich für vier kreisförmige Signaturen sieben unterscheidbare Anordnungen konstruieren, die in Abbildung 15 zu sehen sind.

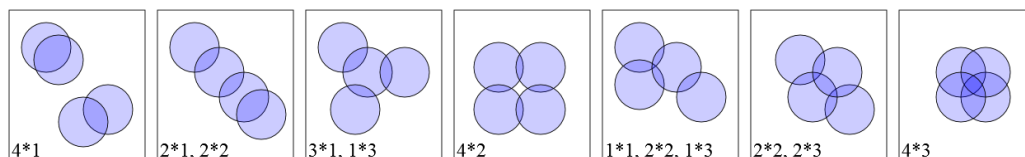


Abb. 15: Topologisch differenzierbare überlappende Anordnungen von 4 transparenten Kreisen

Für jede Konstellation ist angegeben, wie viele Objekte sich mit wie vielen anderen überlappen. In der ersten Konstellation von links z.B. überlappt jede der vier Signaturen mit einer anderen ($4*1$), bei der zweiten zwei Signaturen jeweils mit einer ($2*1$) und zwei Signaturen mit zwei ($2*2$). Es lässt sich bei Berücksichtigung der Anzahl der pro Signatur überlappenden Signaturen ein Qualitätsmaß definieren, dessen Wert von links nach rechts abnimmt, indem die Summe der Produkte gebildet wird und bei gleichem Wert nach dem größten rechten Faktor sortiert wird. Eine Konstellation mit einer besseren Bewertung durch dieses Qualitätsmaß bedeutet jedoch nicht unbedingt eine bessere Lesbarkeit der Signaturen, wie in Abbildung 16 zu sehen ist. So ist beispielsweise die Haltestellensignatur bei Anordnung (a) schlechter zu erkennen als jede Signatur bei (d) und (e). Insbesondere bei komplexen nicht-transparenten Signaturen ist die Anordnung und die Lage der Überlappungsfläche relevanter für die Lesbarkeit als die Anzahl der an der Überlappung beteiligten Objekte.



Abb. 16: Topologisch differenzierbare überlappende Anordnungen von 4 Signaturen

Im Gegensatz dazu ist bei der Zählweise, bei der nur die Anzahl der beteiligten Objekte berücksichtigt wird, der Wert aller Fälle in Abbildung 16 jeweils 4. Jedoch bedeutet ein kleinerer Wert immer, dass mehr Objekte garantiert lesbar sind, so

dass an Hand dieses Maßes ein sinnvoller Rückschluss auf die Lesbarkeit möglich ist. Aus diesem Grund wird als Kriterium für die Darstellbarkeit der Punktmengen die Anzahl der Objekte verwendet, die an einer nicht auflösbaren Überlappung beteiligt sind.

Für die folgenden Betrachtungen zur Lösbarkeit von Überlappungen, wird vorausgesetzt, dass die Punktsignaturen gleich groß sind und dass die maximale Distanz, um die eine Signatur insgesamt verschoben werden darf, gleich dem Radius der Signaturen ist. An einer nicht auflösbaren Überlappung sind folglich mindestens drei Signaturen beteiligt, da auf Grund der gewählten maximalen Verschiebungsdistanz eine Überlappung von nur zwei Signaturen immer lösbar ist. Bei der Abschätzung, ob eine Überlappung lösbar ist betrachten wir zunächst die Dreiecke, die bei einer Delaunay-Triangulation der Punktmenge entstehen. Durch die Verwendung der Delaunay-Triangulation ist sichergestellt, dass alle Objekte, die an einer nicht auflösbaren Dreier-Konstellation beteiligt sind, gefunden werden. Wie in Abbildung 17 dargestellt kann bei einer beliebigen Triangulation (linke Seite) der Fall auftreten, dass drei Punkte, deren Signaturen sich überlappen, nicht durch ein Dreieck verbunden sind. Bei einer Delaunay-Triangulation (Abbildung 17, rechte Seite) ist dieser Fall durch die Umkreisbedingung ausgeschlossen.

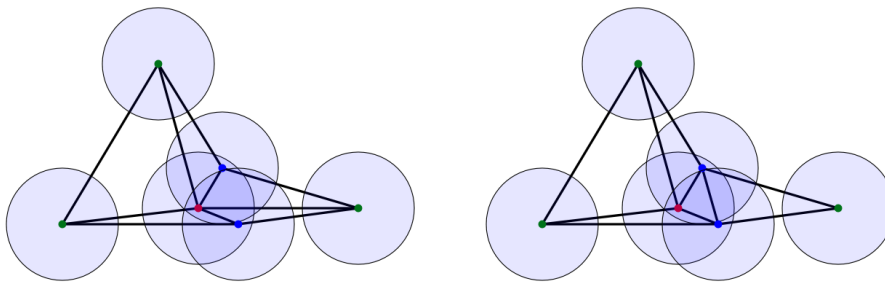


Abb. 17: Detektion von Überlappungen mit beliebiger und Delaunay-Triangulation

Die dichteste Kreispackung einer großen Anzahl von Kreisen auf einer unbegrenzten Fläche entsteht durch eine hexagonale Anordnung der Mittelpunkte [Ne82], d.h. für drei Punkte dass sie ein gleichseitiges Dreieck bilden, dessen Kantenlänge dem Kreisdurchmesser entspricht. Wenn eine Überlappung auflösbar ist, dann ist es möglich, eine Anordnung, die in Abbildung 18 dargestellt ist, zu konstruieren, so dass in jedem der Kreise genau ein Mittelpunkt der sich überlappenden Signaturen liegt. Kann eine solche Anordnung nicht gefunden werden, ist die Überlappung nicht lösbar. Eine effiziente Entscheidung, ob eine solche Anordnung existiert, ist nicht bekannt. Jedoch kann eine Abschätzung, ob eine entsprechende Platzierung möglich ist, durch Betrachtung der Soddy-Kreise vorgenommen werden, die in Abbildung 18 in rot dargestellt sind [De07]. Da die betrachteten Kreise extern tangential zueinander sind, existieren in diesem Fall genau zwei Kreise, die jeweils alle drei gegebenen Kreise tangieren und die als innerer und äußerer Soddy-Kreis bezeichnet werden.

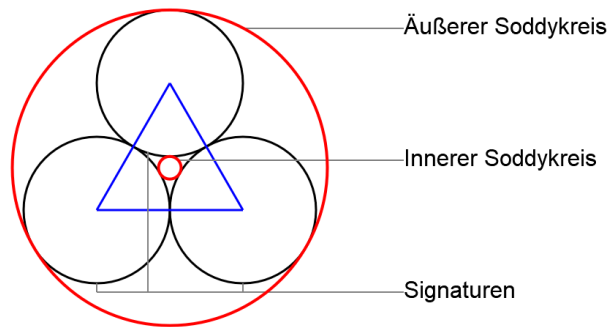


Abb. 18: Innerer und äußerer Soddy-Kreis

Wenn nun die drei Punkte so dicht beieinander liegen, dass ihr Umkreis kleiner als der innere Soddy-Kreis ist, dann ist die Überlappung nicht auflösbar. Um die Anzahl der Objekte zu bestimmen die an nicht auflösbaren Überlappungen beteiligt sind, wird für jedes Delaunay-Dreieck der Umkreis berechnet und mit dem Radius des inneren Soddy-Kreises verglichen. Der Radius des inneren Soddy-Kreises r_i ist lediglich abhängig von den Radien r_1 , r_2 und r_3 der drei gegebenen extern tangentialen Kreise:

$$r_i = \frac{r_1 r_2 r_3}{r_1 r_2 + r_2 r_3 + r_3 r_1 + 2\Delta} \quad \text{mit } \Delta = \sqrt{r_1 r_2 r_3 (r_1 + r_2 + r_3)}$$

Da hier die drei Kreise gleich groß sind, ergibt sich:

$$\begin{aligned} r_i &= \frac{r^3}{3r^2 + 2\Delta} \quad \text{mit } \Delta = r^2\sqrt{3} \\ &= \frac{1}{3 + 2\sqrt{3}} r \approx 0,1547r \end{aligned}$$

Für Konstellationen mit mehr als drei Punkten lassen sich bei der Betrachtung von dichtesten Kreispackungen innerhalb von Kreisen ebenfalls Kriterien finden, unter denen die Überlappungen der Signaturen nicht auflösbar sind. Voraussetzung hierfür ist, dass die dichteste Kreispackung bekannt ist, wie es für einen bis zu dreizehn Kreisen, sowie für neunzehn Kreise der Fall ist ([Pi69], [Me94], [Fo99], [Fo00], [Fo03]). Des Weiteren sollte, analog zur Delaunay-Triangulation für Dreierkonstellationen, eine effiziente Möglichkeit bestehen, alle relevanten Teilmengen zu bestimmen. Dies kann für Vierer-, Siebener- und Neunzehnerkonstellationen folgendermaßen bewerkstelligt werden:

Bei der dichtesten Kreispackung von vier Kreisen sind die Mittelpunkte der Kreise in einem Quadrat mit einer Seitenlänge von $2r$ angeordnet, wie in Abbildung 19 auf der linken Seite zu sehen ist. Wie auch bei der Dreierkonstellation gilt,

dass Überlappungen von vier Signaturen nicht aufgelöst werden können, wenn ihre POIs innerhalb des inneren Kreises liegen. Der Radius des inneren Kreises r_i entspricht der Diagonalen des Quadrats abzüglich des zweifachen Signaturreadius r , also: $r_i = r + r\sqrt{2}$. Die Überprüfung kann ebenfalls mit Hilfe der Delaunay-Triangulation erfolgen, indem die Länge jeder Kante sowie der Abstand der Punkte, die jeweils die beiden Dreiecke zu denen die Kante gehört vervollständigen, berechnet und mit dem Radius r_i verglichen wird. Da es sich bei diesen Strecken um die Diagonalen des Vierecks handelt und die längste Diagonale dem Durchmesser des kleinsten umschließenden Kreises eines Vierecks entspricht, handelt es sich um eine unlösbare Konstellation, falls beide Strecken kürzer als $2r_i$ sind.

Bei der dichtesten Kreispackung von sieben Kreisen sind sechs Kreise extern tangential um den siebten angeordnet (siehe Abbildung 19, Mitte). Daraus folgt, dass sobald sich mehr als sechs andere POIs innerhalb einer Signatur befinden, diese nicht mehr überlappungsfrei dargestellt werden kann.

Die dichteste Kreispackung von neunzehn Kreisen entspricht der Anordnung von sieben Kreisen um die zwölf weitere Kreise platziert werden, wie in Abbildung 19 auf der rechten Seite zu sehen ist. Wäre der innere Kreis genauso groß wie der umschließende Kreis bei der dichtesten Kreispackung von sieben Punkten, so wäre die Existenz einer Signatur, die sich mit 18 weiteren überlappt ein hinreichendes Kriterium, für eine unlösbare Neunzehnerkonstellation. Da der innere Kreis jedoch um $0,09r$ kleiner ist, ist zusätzlich zu überprüfen, dass die Distanz von mindestens sieben Signaturen kleiner als $1,91r$ ist.

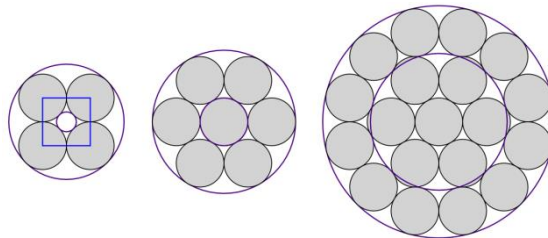


Abb. 19: Dichteste Packungen von 4, 7 und 19 Kreisen in einem Kreis

Entwurf von Signaturen für Sentiments

Die Sentiments, die in dieser Arbeit betrachtet werden, beziehen sich auf POIs, wie z.B. Denkmäler, Schlösser, Restaurants, Parks, Seen, Straßen oder Türme. Die Darstellung dieser Objekte erfolgt abhängig vom Maßstab und der Objektart durch eine Punkt-, Flächen- oder Liniensignatur. Um nun die Sentiments, die sich auf diese Objekte beziehen, zu visualisieren, kann entweder die Signatur des Objektes angepasst werden oder eine zusätzliche Signatur speziell für die Sentiments generiert werden. Welche dieser beiden Varianten sinnvoll ist, hängt sowohl von der ursprünglichen Signatur des Objektes, als auch von weiteren Karteninhalten ab. In Abbildung 20 ist der gleiche Ausschnitt von Braunschweig unter Verwendung unterschiedlicher Signaturen dargestellt. In der rechten Karte wäre es möglich, Sentiments die sich auf Gebäude beziehen durch Änderung der Farbe der Gebäudesignatur zu visualisieren. Bei der linken Darstellung ist dies nicht möglich, da die Farbe der Gebäudesignatur bereits für die Unterscheidung von öffentlichen und nicht-öffentlichen Gebäuden verwendet wird. Eine Visualisierung der Sentiments bezüglich öffentlicher Gebäude wäre in beiden Fällen jedoch durch eine farbliche Anpassung der Punktsignaturen möglich, wobei die gleiche Farbskala für unterschiedliche Objektarten, wie z.B. Museen und Theater verwendet werden könnte. Ein Beispiel für die Darstellung einer zusätzlichen Information durch eine eigene Punktsignatur, die eine vorhandene Punktsignatur ergänzt, ist in Abbildung 20 bei den Parkhäusern zu sehen. Die Auslastung der Parkhäuser wird durch den Anteil der farbigen Fläche an der Gesamtfläche der zusätzlichen Signatur visualisiert, wobei für Auslastungen unter 50% grün und für über 50% rot verwendet wird.



Abb. 20: Einfluss der Hintergrundkarte auf den Signaturentwurf

Bei der Verwendung einer zusätzlichen Signatur für Objekte, die bereits durch eine Punktsignatur dargestellt werden, steigt der Platzbedarf zur Visualisierung der Objekte, so dass bei nachträglichem Einfügen die Lesbarkeit benachbarter Objekte beeinträchtigt werden kann. Somit ist die Verwendung von speziellen Punktsignaturen eher für Objekte, die durch eine Flächensignatur dargestellt werden, oder als Ersatz der bisherigen Punktsignatur geeignet, wobei dann allerdings die Information über die Objektart nicht mehr vorhanden ist. Des Weiteren beeinflussen POIs nicht immer nur die Sentiments des Ortes, an dem sich der zugehörige POI befindet, sondern auch die der Orte, von denen aus der POI zu sehen ist. Dies betrifft insbesondere große Bauwerke. Daher können Sentiments auch als Kontinua aufgefasst und entsprechend visualisiert werden.

Welche der drei Darstellungsarten (angepasste Signatur, zusätzliche Signatur, Felddarstellung) am geeignetsten ist, hängt sowohl von den Objektarten der darzustellenden POIs ab, als auch vom Maßstab. Wir betrachten im Folgenden für jede der drei Darstellungsarten ein konkretes Beispiel:

- a) Eine großmaßstäbliche Darstellung von Sentiments von unterschiedlichen Objektarten durch eine Anpassung der Objektsignaturen. Hierbei bleiben die Objektarten für den Benutzer unterscheidbar und die Signaturen sind dem Kartenleser vertraut.
- b) Eine großmaßstäbliche Darstellung von Sentiments als Kontinua mit der Zielsetzung, dass positiv und negativ wahrgenommene Bereiche vom Leser identifiziert werden können. Für jeden Bereich erfolgt eine aggregierte Darstellung der Sentiments aller POIs, die die Beurteilung des Bereichs beeinflussen.
- c) Eine kleinmaßstäbliche Darstellung von Sentiments für eine Objektart durch eine Punktsignatur, die an Stelle der Objektsignatur verwendet wird. Bei dieser Darstellungsart bestehen die geringsten Einschränkungen beim Design der Signaturen.

Im folgenden Abschnitt 6.1 wird zunächst festgelegt, welche Informationen über Sentiments dem Kartenleser vermittelt werden sollen, bevor in Abschnitt 6.2 an Hand eines Beispiels für Flächensignaturen, Möglichkeiten zur Anpassung vorhandener Signaturen diskutiert wird. In Abschnitt 6.3 wird dann eine Kontinua-Darstellung für Sentiments und in Abschnitt 6.4 verschiedene Punktsignaturen vorgestellt, bevor Abschnitt 6.5 das Kapitel mit einem Fazit abschließt.

6.1. Kommunikationsziel

Als Sentimentmodell wird das eindimensionale Modell verwendet, in dem jeweils nur ein Wert zur Darstellung eines Sentiments angegeben wird. Dennoch ist eine Reihe von Entwurfsentscheidungen für die Visualisierung zu treffen, insbesondere, wenn für einen POI mehrere Sentimentwerte dargestellt werden sollen. In Bezug auf den Wertebereich für die Sentiments gibt es drei Optionen:

- a) Nur zwei Werte, d.h. positiv und negativ.
- b) Mehrere disjunkte Klassen, z.B. sehr positiv/negativ, positiv/negativ, leicht positiv/negativ
- c) kontinuierliche Werte von negativ zu positiv mit oder ohne festgelegte Minimal- und Maximalwert

Im Folgendem wird von Option c) ausgegangen, wobei Werte zwischen -1 und 0 für negative Sentiments verwendet werden und Werte zwischen 0 und 1 für positive. Der Grund hierfür ist zum einem, dass es als nicht notwendig angesehen wird exakte Sentimentwerte aus der Darstellung abzulesen. Es genügt wenn der Leser einen realistischen Eindruck der Meinungen der Reviewer erhält. Zum anderen ist es möglich im Nachhinein Visualisierungen, in denen Signaturen für kontinuierliche Werte erstellt wurden, in eine Darstellung diskreter Klassen umzuwandeln.

Aufgrund der Anzahl an Reviews, die zu einem POI existieren, ist es nicht sinnvoll, jedes Review einzeln zu visualisieren, sondern die Sentiments aller Reviews zu einem Ort zuerst zu aggregieren und dann darzustellen. Da aus der Darstellung ersichtlich sein soll, ob ein POI/ein Gebiet nur positiv, oder nur negativ bewertet wird, oder die Meinungen polarisiert sind, werden die positiven und negativen Werte getrennt aggregiert und jeweils das arithmetische Mittel gebildet. Außerdem soll auch die Anzahl der aggregierten Reviews dargestellt werden.

Ziel ist es also, dass der Kartenleser mit Hilfe der Visualisierung in die Lage versetzt wird folgende Fragen zu beantworten:

- Wird ein Gebiet/POI positiver/negativer gewertet als ein anderer/s?
- Ist die Meinung zu einem Gebiet/POI einheitlich?
- In welchen Gebieten befinden sich hauptsächlich positiv/negativ bewertete POIs?
- Für welche Gebiete liegen viele/wenige Bewertungen vor?

Somit ergeben sich folgende drei Bedingungen, die in jeder der drei Darstellungsvarianten erfüllt werden sollten:

- R1 Geordnete Wahrnehmung zwischen positiven und negativen Sentiments.
- R2 Selektive Wahrnehmung der positive bzw. negative Sentiments.
- R3 Geordnete Wahrnehmung der Anzahl der Reviews.

Unabhängig vom Maßstab sind zunächst die Komponenten der darzustellenden Information, also den Sentiments, festzulegen. Da für jedes Review jeweils ein positiver und ein negativer Sentimentwert vorliegt, bietet es sich an, zur Modellierung zwei Komponenten zu verwenden: eine qualitative, die entweder positiv oder negativ sein kann, und eine quantitative, die die Stärke des Sentiments angibt. Die Information über die Anzahl der vorhandenen Reviews kann dann über eine weitere quantitative Komponente visualisiert werden.

6.2. Darstellung von Sentiments durch Anpassung von Signaturen

Bei der Visualisierung von Sentiments durch Anpassung von vorhandenen Signaturen sind die Gestaltungsmöglichkeiten zum einem durch die schon verwendeten graphischen Variablen beschränkt. Zum anderen sollte auch nach der Anpassung der Signaturen eine assoziative Wahrnehmung der Objektarten erhalten bleiben und die Darstellung von Sentiments für alle Objektarten einheitlich erfolgen.

Für eine Flächensignatur sind nur vier graphische Variablen anwendbar: Farbe, Hellwert, Textur und Richtung. Jedoch ist die zusätzliche Verwendung von Punktsignaturen (Muster) zur Visualisierung von Eigenschaften flächenhafter Objekte möglich, sofern Textur und Richtung nicht verwendet werden. Von den vier, auf Flächensignaturen anwendbaren Variablen, sind der Hellwert und die Richtung am besten für quantitative Attribute geeignet und die anderen beiden Variablen eignen sich für nominale Attribute. In vielen Karten, z.B. der TK25, werden jedoch sowohl Muster als auch Farbe für die Unterscheidung der Objektarten verwendet, wie der Ausschnitt der Legende in Abbildung 21 zeigt. In diesem Fall ist es nicht möglich, eine Anpassung der Signaturen durchzuführen, die alle oben genannten Bedingungen erfüllt.

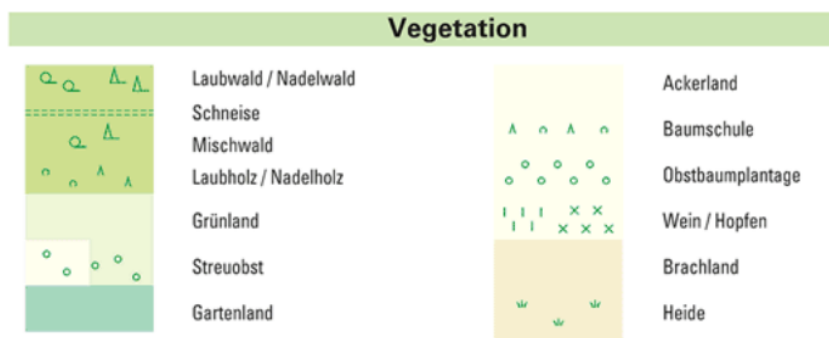


Abb. 21 Ausschnitt der Legende der TK25

Auch bei Karten in denen ausschließlich Muster zur Unterscheidung von Objektarten verwendet werden, ist eine Anpassung von Flächensignaturen zur Darstellung von Sentiments nicht möglich, da nur zwei graphische Variablen verwendet werden können jedoch drei Komponenten zu visualisieren sind. Somit ist eine Visualisierung der Sentiments auf diese Weise nur dann sinnvoll möglich, wenn kein Muster für die Flächensignaturen der Karte verwendet wird.

Ein Muster besteht aus regelmäßig angeordneten Punkt-, Linien- oder Flächensignaturen. Im Folgenden betrachten wir Muster, die aus einzelnen identischen Punktsignaturen zusammengesetzt werden. Für nominale Komponenten sind Farbe, Form und Richtung am besten geeignet. Da jedoch Farbe schon in der Flächensignatur selbst zur Unterscheidung der Objektarten verwendet wird, betrachten wir für die Kodierung der nominalen Komponente der Sentiments nur Form und Richtung. Um die Stärke der Sentiments zu visualisieren, steht nun, neben der Hel-

lichkeit, die Größe der Punktsignatur zur Auswahl. Da es für POIs in der Regel sowohl positive als auch negative Sentiments gibt, ist es wichtig, dass auch die Überlagerung der beiden Muster für den Leser möglichst einfach zu interpretieren ist.

Als konkretes Beispiel werden zwei verschiedene Muster betrachtet, die jeweils aus einer regelmäßigen Anordnung einer Einzelpunktsignatur bestehen, für die in Abbildung 22 mehrere Kombinationen von Sentimentwerten dargestellt sind. Die Stärke der positiven Sentiments nimmt entlang der x-Achse und die Stärke der negativen entlang der y-Achse zu. In beiden Fällen symbolisiert die Größe der Signatur ihre Stärke. Im linken Teil der Abbildung wird die Richtung der Signatur verwendet, um zwischen positiven und negativen Gefühlen zu unterscheiden und im rechten Teil die Form: Für positiv ein Plus und ein Kreis für negativ. Hier wird als Größe der Durchmesser des Kreises verwendet und beim Plus die Länge der Striche, so dass bei gleicher Größe die Eckpunkte des Pluszeichens auf dem Kreis liegen.

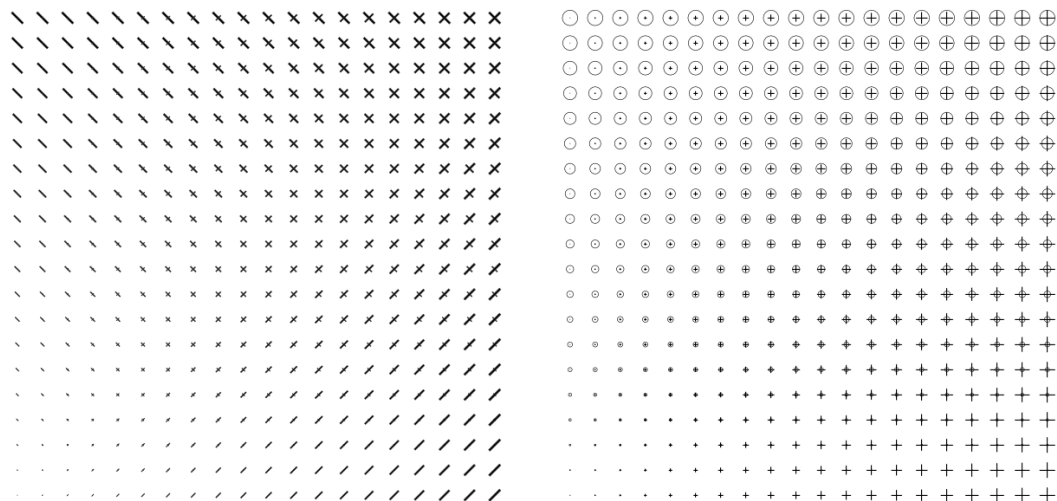


Abb. 22 Verschiedene Wertkombinationen bei Kodierung der Sentimentstärke durch Größe und von “positiv/negativ” durch Richtung (links) und Form (rechts)

In Abbildung 23 wird anstatt der Größe der Hellwert zur Visualisierung der Sentimentstärke verwendet, ansonsten sind die Signaturen analog zu Abbildung 22 konstruiert. Da sowohl die Größe als auch der Hellwert die Sichtbarkeit der Signatur verändern, ist bei beiden Varianten klar zu erkennen, ob die positiven oder die negativen Sentiments überwiegen. Bei der Größe sind jedoch geringe Unterschiede zwischen den positiven und negativen Werten besser wahrnehmbar. Bei beiden Varianten sind die Unterschiede bei größeren Werten deutlicher sichtbar, da bei geringen positiven und negativen Werten die Signatur insgesamt schlechter wahrnehmbar ist. Dies bedeutet, dass durch die Verwendung der Variablen eine Gewichtung der Information vorgenommen wird, die in diesem Anwendungsfall wün-

schenswert ist, da insbesondere Orte mit sehr positiven und/oder sehr negativen Sentimentwerten relevant sind.

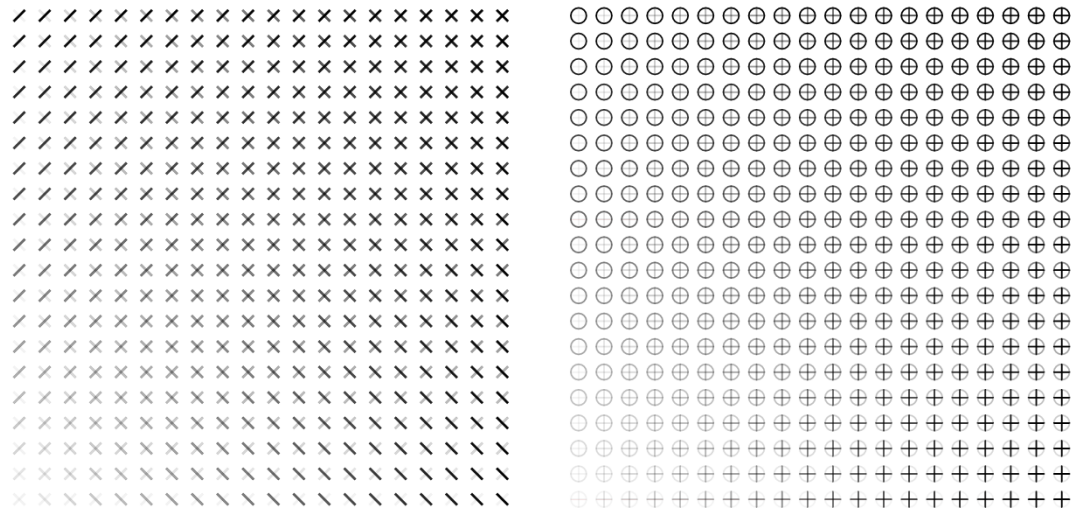


Abb. 23: Verschiedene Wertkombinationen bei Kodierung der Sentimentstärke durch ihren Hellwert und von "positiv/negativ" durch Richtung (links) und Form (rechts)

Für die dritte Komponente, die Anzahl der Reviews, kann, genauso wie für die Sentimentstärke, die Größe oder der Hellwert der Einzelpunktsignatur verwendet werden, so dass vier verschiedene Visualisierungen möglich sind:

Tabelle 2: Kombinationen graphischer Variablen zur Sentimentdarstellung

	Positiv/negativ	Sentimentstärke	Reviewanzahl
a)	Richtung	Größe	Hellwert
b)	Form	Größe	Hellwert
c)	Richtung	Hellwert	Größe
d)	Form	Hellwert	Größe

Hierbei ist zu beachten, dass sich zum einem die Größe und der Hellwert gegenseitig beeinflussen und zum anderen der Hellwert der ursprünglichen Flächensignatur eine Änderung der Wahrnehmung des Hellwerts des Musters bewirken kann. Um diese Effekte zu visualisieren, sind in Abbildung 24 alle vier Varianten dargestellt. Als Hintergrundkarte dient der Lageplan eines Zoos, in dem die Gehege in unterschiedlichen Grüntönen dargestellt werden. Die Sentimentwerte und Reviewanzahlen sind synthetisch erzeugt.

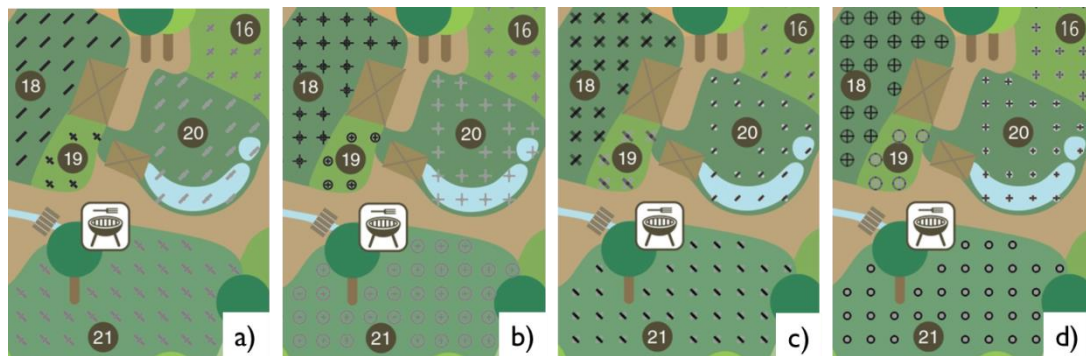


Abb. 24: Varianten zur Anpassung von Flächensignaturen mit unterschiedlichem Hellwert

Die Unterschiede in den Hellwerten sind auffälliger, so dass bei den Varianten a) und b) leichter zu sehen ist, über welche Gebiete weniger Informationen vorliegen. Bei den Varianten c) und d) sind die Hellwerte für positive und negative Signaturen unterschiedlich, dadurch wird eine Anordnung der Flächen auf Grund des Hellwerts erschwert. Für die Flächen 16 und 21 liegen gleich viele Bewertungen vor, für die Fläche 20 etwas weniger. Da die Hellwerte der farbigen Flächensignaturen jedoch unterschiedlich sind, wird eine korrekte Anordnung der Hellwerte des Musters erschwert. In Abbildung 25 sind die farbigen Flächensignaturen so abgeändert, dass sich die Farbwerte stärker unterscheiden, die Hellwerte aller Farben aber gleich sind. Die Sentimentwerte und die Muster sind identisch zu Abbildung 24, sind aber deutlicher zu erkennen und einfacher miteinander zu vergleichen.

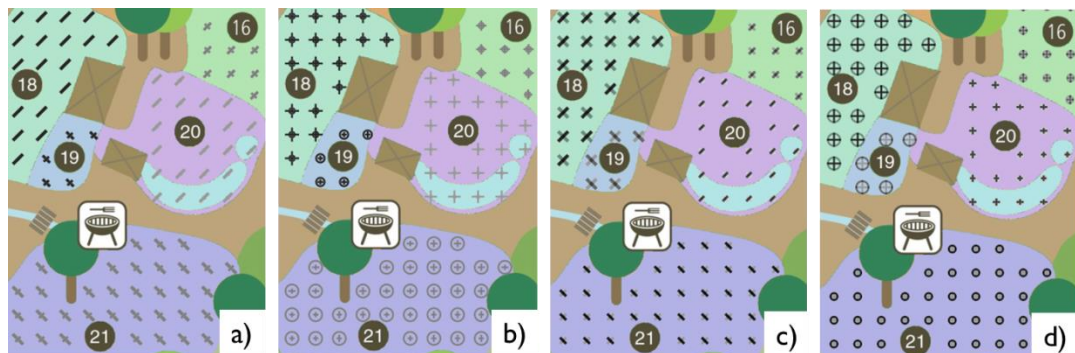


Abb. 25 Varianten zur Anpassung von Flächensignaturen mit gleichem Hellwert

6.3. Darstellung von Sentiments als Kontinua

Die Zielsetzung dieser Darstellungsvariante ist, dass positiv und negativ wahrgenommene Bereiche vom Leser identifiziert werden können und dass erkennbar ist, ob für einen Bereich nur positive Sentiments oder nur negative vorliegen, oder ob

die Meinungen über den Ort polarisiert sind. Die Bereiche sind hierbei nicht a priori vorgegeben, sondern ergeben sich aus den Einflussbereichen der POIs. Anders als bei der Darstellungsvariante in Abschnitt 6.2 kann für die Flächensignatur kein Muster verwendet werden. Dies liegt daran, dass sich beliebig viele Einflussbereiche überlappen können und eine Überlagerung verschiedener Muster unlesbar wird. Gleiches gilt für Textur und Orientierung. Folglich bleibt nur die graphische Variable Farbe für die nominale Komponente (positiv oder negativ) und der Tonwert (Helligkeit) für die Stärke der Sentiments (quantitative Komponente). Die Darstellung einer dritten Komponente (Anzahl der Reviews) ist bei dieser Variante nicht möglich.

Eine intuitive Wahl für Farben, um positive und negative Werte auszudrücken, ist grün und rot (Ampel-Metapher). Die Veränderung des Tonwertes eines Objekts hat theoretisch die gleiche visuelle Wirkung wie die Veränderung seiner Transparenz, solange der Hintergrund weiß ist. Die Transparenz bietet jedoch Vorteile bei der Erstellung von Karten. Zum einem ist die Basiskarte noch sichtbar, was für den Benutzer notwendig ist, um die Informationen auf der Karte mit Orten in der realen Welt zu verknüpfen. Zum anderen müssen die Werte für Bereiche, deren Beurteilungen von mehr als einem POI beeinflusst werden, nicht explizit berechnet werden, da die Überlagerung von transparenten Farben automatisch zu einer anderen Farbe führt, die für diese Kombination von Werten steht. Als Form für die Signaturen der Sentiments werden Kreisscheiben verwendet, deren Zentrum der Mittelpunkt des POIs ist, auf den sie sich beziehen, und deren Radius von ihrer Fläche abhängt. Darüber hinaus nehmen wir an, dass der Einfluss eines POIs mit zunehmender Distanz abnimmt, so dass die Transparenz der Scheiben mit zunehmender Entfernung zu ihrem Zentrum zunimmt. Abbildung 26 zeigt 20 Kreisscheiben, die positive Werte von 0 bis 1 in Schritten von 0,05 auf der oberen Hälfte und den entsprechenden negativen Werten in der unteren Hälfte darstellen.

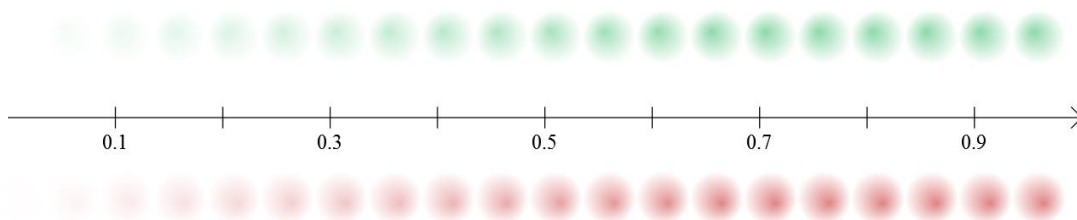


Abb. 26: Positive und negative Sentimentwerte

Da jeder POI zwei Werte (positiv und negativ) erhalten kann, sollte die Deckkraft der Scheiben nicht 100 Prozent erreichen. Für verschiedene Schattierungen von Rot und Grün wurde eine Matrix von Kombinationen von roten und grünen Werten aufgetragen, um festzustellen, ob die Kombination von Farben auch zu intuitiven Ergebnissen führt. Abbildung 27 zeigt die Matrix für die Farben, die im

Folgenden verwendet wird. Um eine bessere Interpretation der einzelnen Werte zu ermöglichen, ist der dunkelste Punkt einer Kreisscheibe nicht das Zentrum, sondern ein wenig nach rechts unten für negative Werte und nach oben links für positive verschoben.

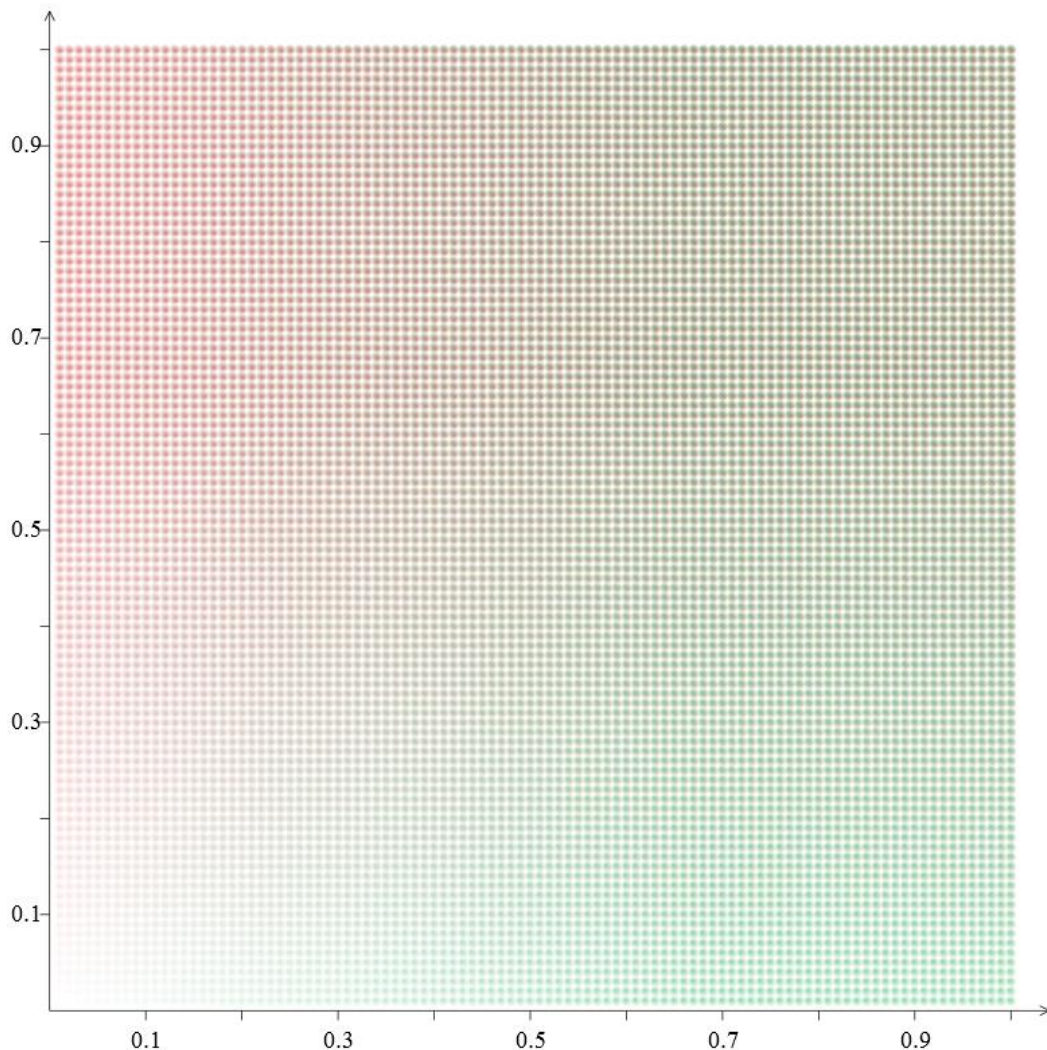


Abb. 27: Kombination positiver und negativer Sentimentwerte

Jedoch ist ein rot-grünes Farbschema nicht für Menschen mit Rot-Grün-Sehschwäche oder für Schwarz-Weiß-Drucke geeignet. Für diese Fälle ist eine einfache Abwandlung der oben genannten Prinzipien möglich um alternative Visualisierungen zu generieren. Eine Möglichkeit ist die Verwendung von schwarz für negative Werte und einer zweiten beliebige Farbe zur Darstellung positiver Werte, z. B. blau und schwarz, wie in Abbildung 28 in der oberen Reihe rechts dargestellt. In Graustufenbildern gibt es nur zwei Farben: weiß für positiv und schwarz für negativ. In diesem Fall sollte der Hintergrund der Karte hellgrau eingefärbt werden, da sonst positive Bereiche schwer zu lokalisieren sind (obere linke Reihe in Abbil-

dung 28). Grundsätzlich kann auch für positive Werte Weiß beibehalten werden und Schwarz durch eine beliebige andere Farbe ersetzt werden, wie z.B. rot (untere linke Reihe in Abbildung 28). Bei der Kombinationen von anderen Farben ist es schwieriger, ein passendes Paar von Farben zu bestimmen, bei dem jede Kombination zu einem passendem und interpretierbaren Ergebnis führt.

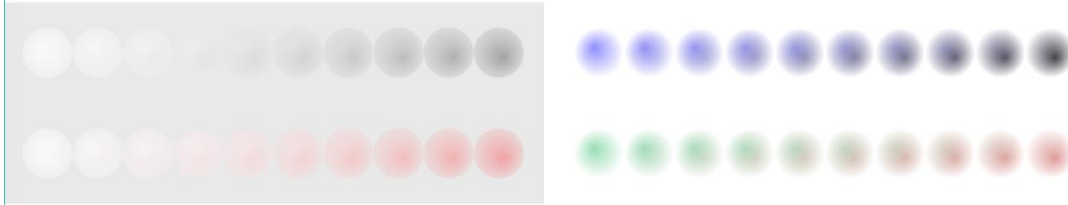


Abb. 28: Verschiedene Farbschemata für Sentiments

Als Beispiel einer kartenähnlichen Graphik mit einer Darstellung von Sentiments als Kontinua sind in Abbildung 29 die Sentimentwerte für 11 Sehenswürdigkeiten in Reykjavik (Hauptstadt von Island) dargestellt. Als Hintergrundkarte dient eine reduzierte Darstellung von OSM-Daten in denen Straßen, Gebäude- und Grünflächenumrisse in Schwarz und Gewässer in Blau gezeichnet sind. Die Sentimentwerte sind aus Virtual-Tourist-Reviews extrahiert.



Abb. 29: Sentimentkarte für Reykjavik

Beispielsweise ist in der südöstlichen Ecke die berühmte Kirche Hallgrímskirkja zu sehen, die positiv bewertet ist. Nördlich des Tjörnin Sees sind vier sich überlappende Sentimentsignaturen, drei positive: für ein Denkmal des Jón Sigurdsson, für das Alþingishúsið (isländische Parlamentsgebäude) und für den Alþingisgarðurin (Parlamentsgarten) und eine leicht negative für die Dómkirkjan (Kirche). Der Ein-

flussbereich der Sehenswürdigkeiten ist in dieser Abbildung durch ihre Fläche und ihre Höhe abgeschätzt worden.

6.4. Punktsignaturen für kleinmaßstäbige Karten

Beim Entwurf von Punktsignaturen für Sentiments bestehen mehr Freiheiten als bei den vorhergehenden Varianten, da alle sechs graphischen Variablen verwendet werden können. Für die nominale Komponente ist Farbe am besten geeignet, um eine selektive Wahrnehmung zu unterstützen. Für die beiden quantitativen Komponenten (Sentimentstärke und Reviewanzahl) könnte neben Hellwert und Größe auch die Richtung verwendet werden, jedoch ist die Bestimmung der Rangfolge bei Variationen der Richtung weniger intuitiv.

Da für Orte sowohl positive als auch negative Sentiments vorliegen, ist zu beachten, dass beide Signaturen für alle Kombinationen von Sentimentwerten und Reviewanzahlen lesbar sein sollten, wenn sie demselben Ort zugewiesen werden. Aus diesem Grund bietet es sich an, für die nominale Komponente neben der Farbe auch die Form oder die Richtung zu variieren. Des Weiteren ist es sinnvoll entweder möglichst einfache Formen zu verwenden, da diese auch in kleineren Größen lesbar sind, und/oder Formen zu verwenden, die Assoziationen mit der Thematik der Komponente wecken. In Abbildung 30 sind sechs verschiedene Punktsignaturenpaare (a) bis f)) dargestellt, die für Sentiments verwendet werden können, die im Folgendem zunächst erläutert und dann in Abschnitt 8.1 evaluiert werden. Für jedes Signaturenpaar sind die positive und die negative Signatur jeweils für die Größenwerte 0.1, 0.5 und 1 und den neun möglichen Kombinationen abgebildet. Der Hellwert ist für die Signaturen (c),(e) und (f) 0.8, dieser Wert wird wie bei der Darstellung als Kontinua (Abschnitt 6.3) als Maximalwert festgelegt, damit auch bei Überlagerung von Signaturen erkennbar ist, dass mehrere Signaturen beteiligt sind.

Bei den Signaturen a) und b) handelt es sich um die Einzelpunktsignaturen die in Abschnitt 6.2 als Muster verwendet wurden mit dem Unterschied, dass zusätzlich Farbe als Variable verwendet wird. Bei c) wird sowohl für positive als auch für negative Sentiments ein Kreis verwendet, so dass die Signatur mit der geringeren Stärke vollständig überdeckt ist, und nur wegen der Transparenz erkennbar ist. Bei d) wird wie bei a) die gleiche Form (ein Halbkreis), aber eine unterschiedliche Richtung verwendet. Als Ankerpunkt wird der Mittelpunkt des vollständigen Kreises verwendet und der Unterschied der Richtung beträgt 180° , so dass sich die beiden Signaturen nicht überlappen.

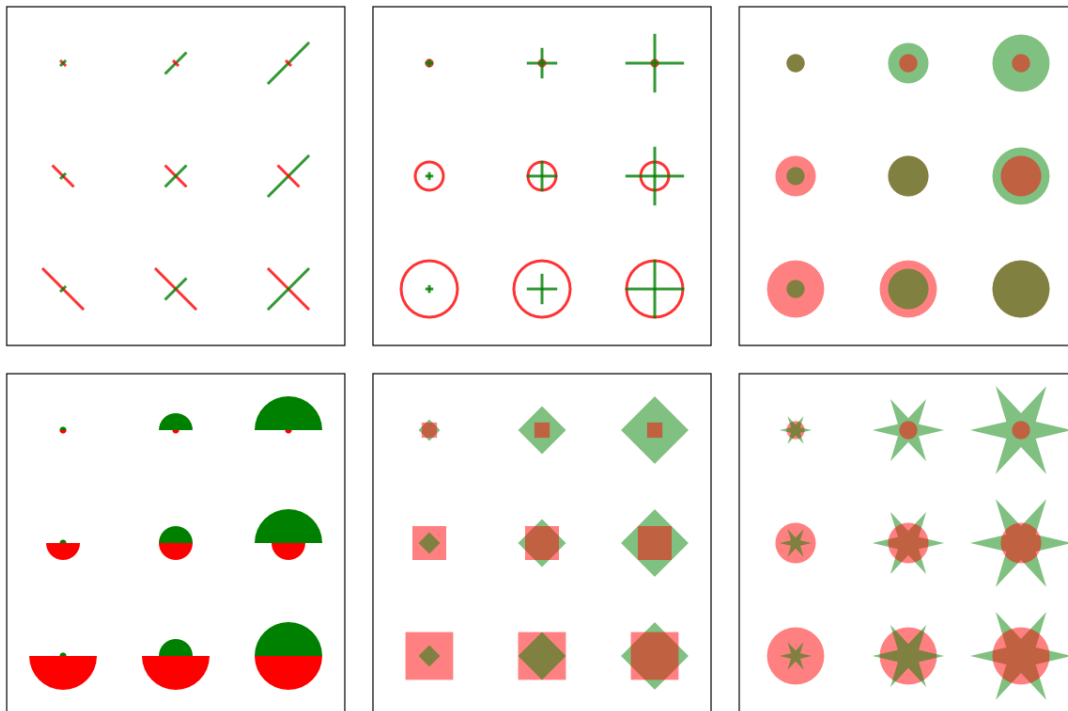


Abb. 30: Kombinierte Punktsignaturen für unterschiedliche Sentimentwerte

Als Größe der Halbkreise wird in Abbildung 30 der Radius verwendet, es ist genauso gut möglich die Fläche zu nehmen. Der Vorteil beim Radius ist das lineare Größenverhältnisse von Menschen genauer abgeschätzt werden können, der Nachteil ist das die Signaturen mit maximalen Sentimentwerten mehr Platz benötigen. Das Signaturenpaar e) besteht aus zwei Quadraten, von denen das positive um 45° gedreht ist. Auch bei diesen Paar kann die kleinere Signatur komplett überdeckt werden. Wenn als Größe der Signatur ihre Fläche aufgefasst wird, tritt dieser Fall ein, falls eine Signatur weniger als halb so groß wie die andere ist. Bei f) wird für negative Sentiments erneut ein Kreis verwendet und für positive Sentiments ein Stern. Sterne werden häufig für Bewertungen verwendet, wobei es sich zumeist um Sterne mit fünf Zacken handelt. Da anders als bei den meisten Bewertungen nur ein Stern verwendet wird, dessen Größe bzw. Hellwert variiert um verschiedene Bewertungen auszudrücken, wird hier ein sechszackiger Stern verwendet. Die Konstruktion des Sterns erfolgt mit Hilfe von zwei ineinander geschachtelten Sechsecken, wobei das innere Sechseck um 15° gedreht ist. Die Stützpunkte der beiden Sechsecke sind auch die Stützpunkten des Sterns und können in Abhängigkeit von den Radien der Umkreise der beiden Sechsecke r_i und r_o berechnet werden, wie in Abbildung 31 zu sehen ist.

Die Fläche des Sterns besteht aus sechs gleichseitigen Dreiecken, die das innere Sechseck bilden und den sechs dreieckigen Zacken des Sterns. Die Formel für die Berechnung der Fläche lautet:

$$A_{star} = 6 \left(\frac{\sqrt{3}}{4} r_i^2 + \left(r_o - \frac{\sqrt{3}}{2} r_i \right) r_i \right)$$

Der Radius der negativen Kreissignatur wird durch gleichsetzen der Flächen der beiden Signaturen berechnet.

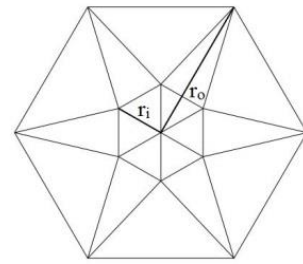


Abb. 31: Konstruktion der Sternsignatur

6.5. Fazit

Bei der Anpassung von Signaturen um zusätzliche Informationen darzustellen sind neben den bereits verwendeten graphischen Variablen auch Wechselwirkungen zwischen ihnen und den neu hinzugefügten Variablen relevant. Hierbei sind nicht nur die Signaturen zu berücksichtigen, die angepasst werden sollen, da beispielsweise das Hinzufügen eines Musters zu einer einfarbigen Flächensignatur auch die Lesbarkeit von komplexen Punktsignaturen beeinflussen kann, die sich auf oder in der Nähe dieser Fläche befinden. Für die Realisierung eines Systems, mittels dem der Benutzer zusätzliche Informationen durch Anpassen von vorhandenen Signaturen visualisieren kann, sind also neben den Informationen über alle verwendeten Signaturen, der Zugriff auf die anzupassenden Signaturen, sowie die Überprüfung möglicher Wechselwirkungen notwendig. Der größte Vorteil dieser Variante ist, dass keine Generalisierung notwendig ist, da der Karte keine Objekte hinzugefügt werden.

Eine Herausforderung bei der Darstellung von zusätzlichen Informationen als Kontinua ist die Abbildung des Einflussbereiches eines POIs auf den ihn umgebenden Raum. Für den Fall das Einflussbereiche sich überlappen können, ist entweder ein geeignetes Interpolationsverfahren zu wählen, oder die Möglichkeit zur Darstellung der Zusatzinformation ist auf transparente farbige Flächen beschränkt. Eine Verfeinerung des in Abschnitt 6.3 vorgestellten Ansatzes wäre es, den Einflussbereich einer Sehenswürdigkeit an Hand ihrer tatsächlichen Sichtbarkeit z.B. mittels Ray Tracings auf Grundlage eines 3D-Stadtmodell zu ermitteln und die berechnete Fläche mit den radialen Signaturen zu verschneiden. Für die Umsetzung eines Systems für die Darstellung als Kontinua bietet sich die Verwendung von Web Map Services (WMSs) [OGC02] an, die vom Open GeoSpatial Consortium (OGC) definiert wurden. Hierzu ist ein WMS zu implementieren, der die Sentimentwerte darstellt und die Hintergrundkarte kann von einem anderen WMS angefragt werden, da Kartenlayer von verschiedenen Anbietern einfach kombiniert werden können, d.h. die Interoperabilität garantiert ist. Jedoch ist zu überprüfen, ob eine Kombination der Layer zu einer lesbaren Karte führt. Damit Farbe als gra-

phische Variable zur Darstellung von Sentiments zur Verfügung steht, ist eine Schwarz-weiß Karte mit einem geringen Schwarzanteil als Hintergrund ideal.

Die Darstellung von zusätzlichen Informationen durch Punktsignaturen bietet, die meisten Wahlmöglichkeiten bei dem Entwurf der Signaturen. Des Weiteren ist diese Darstellungsart am weitesten verbreitet, da sie in prominenten existierenden Systemen, wie z.B. Google Maps verwendet wird. Allerdings ist es für die Lesbarkeit der Karte wichtig, dass sich die Symbole nicht überlappen, insbesondere wenn die verwendeten Signaturen nicht identisch sind, sondern über ihre graphischen Variablen weitere Informationen visualisiert werden sollen. Aus diesem Grund sind, für die Realisierung eines Systems, das den Benutzer bei der Generierung lesbarer Karten unterstützt, Generalisierungsverfahren zur Auflösung von Überlappungen von Punktsignaturen notwendig.

Verdrängungsverfahren für Punktsignaturen

Die grundlegende Idee des in dieser Arbeit entwickelten Verfahren ist es, jedem Punkt eine eindeutige Fläche zuzuordnen, die groß genug ist, um die Signatur komplett zu enthalten. Um dieses zu erreichen, wird eine iterative Methode entwickelt, die für gleichgroße Signaturen ein Voronoi-Diagramm und für unterschiedlich große Signaturen ein Power-Diagramm als Hilfsstruktur verwendet. Für das Verfahren sind zwei Eingabeparameter zu setzen: die Größe der Punktsignaturen und die maximal zulässige Verschiebungsdistanz. Der Fokus der in diesem Kapitel entwickelten Verfahren liegt auf der Verdrängung der Punktsignaturen gegeneinander, da es sich bei ihnen um die wichtigste Information (dem Vordergrund), handelt. Eine Verdeckung der Hintergrundobjekte ist in den meisten Fällen, wie z.B. der bebauten Fläche durch die Punktsignatur in Abbildung 32, unproblematisch. Im Gegensatz dazu kann eine Änderung der Topologie bzw. der relativen Lage zwischen Vorder- und Hintergrundobjekt zu Irritationen des Kartenlesers führen. Deswegen soll das Verfahren die Möglichkeit bieten, einzelne Hintergrundobjekte, die zur räumlichen Einordnung der Punktsignaturen relevant sind, wie z.B. die Autobahnen und Flüsse in Abbildung 32, zu berücksichtigen.

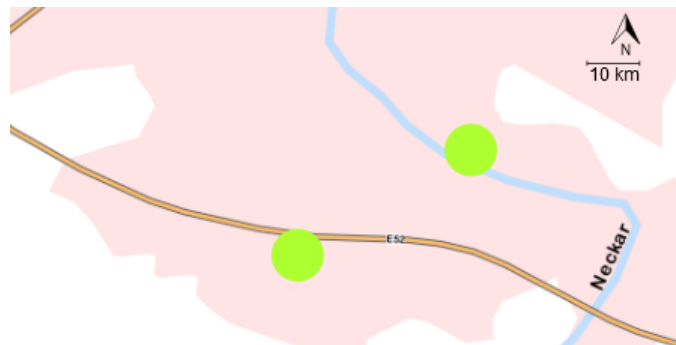


Abb. 32: Überdeckung relevanter Hintergrundobjekte durch Signaturen

In Abschnitt 7.1 wird zunächst das Verfahren für gleich große kreisförmige Signaturen entwickelt. Für die Bestimmung der Verschiebungsrichtung werden zwei Heuristiken (Abschnitt 7.1.1) vorgestellt. Die Berücksichtigung von Hintergrundobjekten und der maximalen Verschiebungsdistanz wird durch eine Begrenzung der Voronoizellen realisiert (Abschnitt 7.1.2). Anschließend wird eine Möglichkeit aufgezeigt die Punktmenge zu zerlegen um die Laufzeit des Algorithmus zu verbessern

(Abschnitt 7.1.3) und mögliche Strategien zur Auswahl der Parameter diskutiert (Abschnitt 7.1.4). In Abschnitt 7.2 werden abschließend Varianten des Verfahrens für andere Formen und unterschiedlich große Signaturen vorgestellt.

7.1. Verdrängung gleich großer kreisförmiger Signaturen

Gegeben sei eine Menge geokodierter POIs und ihre kreisförmige Signatur fixer Größe für einen definierten Maßstab. Anfangs wird das Voronoi-Diagramm der Koordinaten berechnet, jedem POI die Fläche seiner Voronoizelle zugeordnet und der Mittelpunkt der Signatur auf die Koordinaten des POIs platziert. In Abbildung 33 werden die drei verschiedenen Fälle gezeigt, die zu berücksichtigen sind: Erstens, die Signatur passt in die Voronoizelle. Zweitens, die Signatur überlappt sich mit ihrem nächsten Nachbarn, aber die Voronoizelle ist groß genug, um die Signatur komplett zu enthalten. Drittens, die Voronoizelle ist zu klein, um die Signatur zu überdecken. In den beiden letzten Fällen müssen die Signaturen verschoben werden und das Voronoi-Diagramm wird angepasst. Anschließend werden die neuen Positionen evaluiert.

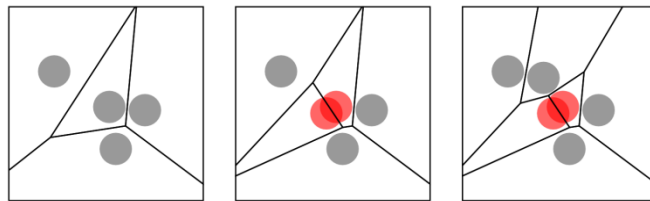


Abb. 33: Anfangsplatzierungen von POIs in Voronoizellen: Mögliche Fälle

Die Methode terminiert wenn es keine Überlappungen mehr zwischen Signaturen gibt, oder wenn kein Punkt mehr verschoben werden kann, ohne den Schwellwert (Radius der Signatur) zu überschreiten, oder eine festgelegte Anzahl an Iterationen überschritten wurde. Der Algorithmus besteht im Prinzip aus zwei ineinander geschachtelten Schleifen (siehe Algorithmus 1). In der inneren Schleife werden die neuen Positionen aller überlappenden Signaturen berechnet und in der äußeren Schleife wird das Voronoi-Diagramm neu berechnet.

```

1: procedure displacePOIs(point set P, radius r)
2:    $P' \leftarrow \emptyset$ 
3:    $P'' \leftarrow P$ 
4:   iter  $\leftarrow 0$ 
5:   while ( $P' \neq P''$  and iter < maxiter) do
6:     iter  $\leftarrow$  iter+1
6:      $P' \leftarrow P''$ 
7:      $P'' \leftarrow \emptyset$ 
8:     vd  $\leftarrow$  calculate Voronoi diagram of  $P'$ 

```



```

9:      for (each  $p \in P'$ ) do
10:         if (symbol( $p$ ) inside vd.cell( $p$ )) do
11:             $P'' \leftarrow P'' \cup \{p\}$ 
12:         else
13:             $P'' \leftarrow P'' \cup \{\text{displace}(p, r)\}$ 
14:         end if
15:      end for
16: end while

```

Algorithmus 1: Prinzip der Punktverdrängung

7.1.1 Heuristiken zur Bestimmung der Verschiebungsrichtung

Die Qualität der Verdrängung hängt vollständig von der Berechnung der Verschiebungsvektoren ab. Um eine gegebene Platzierung zu verbessern, sollten kleine Voronoizellen (Fall 3) vergrößert werden und die Orte sollten näher am Mittelpunkt ihrer Zelle liegen (Fall 2). Um dieses zu erreichen, verwenden wir zwei alternative Heuristiken:

- a) Verschiebung des Punktes in Richtung des Mittelpunkts des größten inneren Kreises seiner Voronoizelle (siehe Abbildung 34). Hierbei handelt es sich um einen Greedy-Ansatz, der gewährleistet, dass die Sichtbarkeit der am wenigsten sichtbarsten Signatur in jeder Iteration erhöht wird.

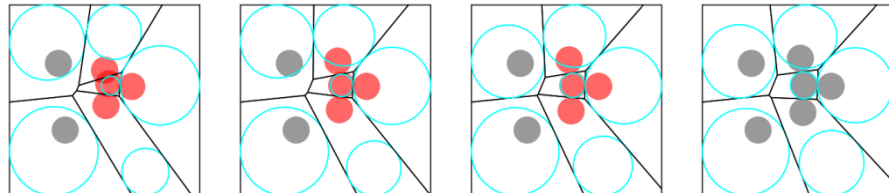


Abb. 34: Iteration 1 bis 3 und Ergebnis durch Heuristik a)

- b) Verschiebung des Punktes in Richtung des Schwerpunktes seiner Voronoizelle (siehe Abbildung 35). Bei diesem Ansatz kann die Sichtbarkeit verringert werden, allerdings können dadurch lokale Optima überwunden werden.

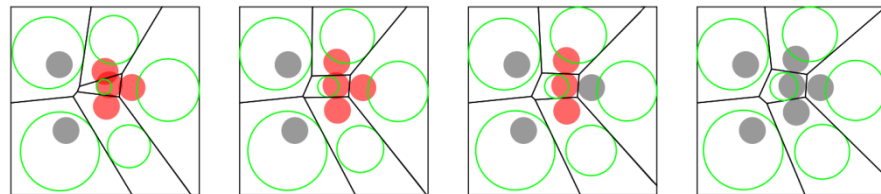


Abb. 35: Iteration 1 bis 3 und Ergebnis durch Heuristik b)

Größter innenliegender Kreis (GIK-Heuristik)

Der Mittelpunkt eines größten innenliegenden Kreises (GIK) eines konvexen Polygons liegt auf einem Knotenpunkt der medialen Achse des Polygons, die in $O(n)$ berechnet werden kann [CSW99]. Sind zwei Kanten des Polygons parallel, können mehrere größte Kreise existieren, von denen für die Heuristik der Kreis gewählt wird, der dem POI am nächsten liegt. Bei Betrachtung des größten gerade noch in die Voronoizelle passenden Kreise, deren Mittelpunkte auf der Strecke vom POI zum Mittelpunkt des GIK liegen, so sind ihre Radien monoton steigend. Dies hat zur Folge, dass die Signaturfläche, die aus der Voronoizelle herausragt, kleiner wird, wenn der Mittelpunkt der Signatur in Richtung des Mittelpunkts des GIK verschoben wird. Ist der größte innenliegende Kreis größer als die Signatur (Fall 2), entstehen durch die Verschiebung des POIs keine neuen Überlappungen und wenn der Abstand zwischen dem Mittelpunkt des GIK geringer ist als der Schwellwert (maximale Verschiebungsdistanz), führt die Heuristik stets zu einer überlappungsfreien Lösung.

Ist der GIK kleiner als die Signatur (Fall 3), können durch die Heuristik neue Überlappungen entstehen, jedoch wird die größte Überlappungsfläche reduziert. Dies liegt darin begründet, dass zum einen der Abstand zum nächsten Nachbarn durch die Verschiebung vergrößert wird und dass zum anderen die Überlappungsfläche zweier Kreise durch Reduktion des Abstandes der Mittelpunkte stärker wächst als sie durch eine Vergrößerung des Abstandes um den gleichen Wert schrumpft. Im folgendem wird zunächst geometrisch gezeigt, dass sich der Abstand zum nächsten Nachbarn eines Punktes durch Anwendung der Heuristik vergrößert. Zur Verdeutlichung der geometrischen Zusammenhänge betrachten wir beispielhaft einen Punkt und seine fünf nächsten Nachbarn, die in Abbildung 36 in blau bzw. in schwarz dargestellt sind. Die Abbildung und die konkrete Konfiguration des Beispiels dienen lediglich zur Veranschaulichung und zur Zuordnung von Farben zu Komponenten, um eine kompaktere Beschreibung zu ermöglichen. Auf der linken Seite ist die Situation vor der Verschiebung zu sehen. Zusätzlich zu den Punkten sind ihre Signaturen sowie die Voronoizelle des blauen Punktes und die Abstände zwischen dem blauen Punkt und seinen Nachbarn eingezeichnet. Auf der rechten Seite ist die Signatur des blauen Punktes auf den Mittelpunkt des GIK der Voronoizelle, der in türkis abgebildet ist, verschoben und die Abstände zum Mittelpunkt der verschobenen Signatur und den Nachbarn des blauen Punktes sind in grün dargestellt. Die Abstände des blauen Punktes zu seinen Nachbarn sind auf der rechten Seite ebenfalls abgebildet, um einen Vergleich der Linienlängen zu erleichtern. Für die folgende Betrachtung werden die Abstände jeweils durch die Voronoikante, die durch den Nachbarn und den blauen Punkt definiert wird, in einen inneren und einen äußeren Abschnitt unterteilt. Beim Vergleich der Länge der Abschnitte zwischen den schwarzen Punkten und dem blauen Punkt (schwarze Linien) und dem Mittelpunkt des GIK (grüne Linien) lassen sich folgende Zusammenhänge erkennen, die unabhängig von der konkreten Anordnung der Punkte gelten:

- Die Länge des äußeren Abschnittes ist bei den schwarzen Linien immer kürzer als die der grünen, da es sich bei den schwarzen Linien um das Lot und damit um die kürzeste Verbindung von dem Punkt auf die Voronoikante handelt.
- Die Länge der inneren Abschnitte der grünen Linien ist immer größer oder gleich dem Radius des GIK, da der Startpunkt der Linien der Mittelpunkt des GIK ist und der GIK komplett innerhalb der Voronoizelle liegt.
- Die Länge der inneren Linienabschnitte ist immer gleich der Hälfte der Gesamtlänge der schwarzen Linie, auf Grund der Definition des Voronoi-Diagramms.
- Die Länge des inneren Linienabschnittes der kürzesten schwarzen Linie ist kleiner gleich dem Radius des GIK, da der Mittelpunkt des GIK der Punkt ist, bei dem der kleinste Abstand zum Rand maximal ist.

Dies bedeutet zusammengefasst, dass sich die kürzeste schwarze Linie aus dem kürzesten inneren und dem kürzesten äußeren aller Liniensegmente zusammensetzt. Daraus folgt, dass der Abstand zwischen dem Mittelpunkt des GIK und seinem nächsten Nachbarn in jedem Fall größer gleich dem Abstand zwischen dem blauen Punkt und seinem nächsten Nachbarn ist.

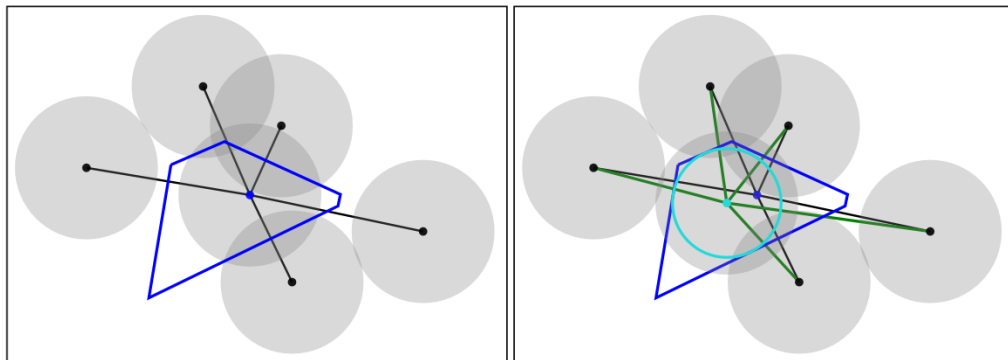


Abb. 36: Vergrößerung des Abstands zum nächsten Nachbarn durch Verschiebung in Richtung des Mittelpunkts des GIK

Im Worst Case verkleinert sich der Abstand zwischen einem anderen Nachbarn um den gleichen Wert, um den sich der Abstand zum nächsten Nachbarn verringert. Dies ist dann der Fall, wenn die Mittelpunkte des nächsten Nachbarn, des GIK und eines anderen Nachbarn auf einer Linie liegen, wie in Abbildung 37 zu sehen.

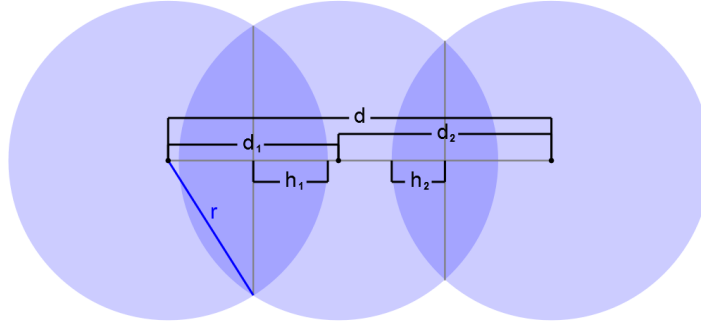


Abb. 37: Änderung der Gesamtüberlappungsfläche der Signaturen

Für den Nachweis, dass mit einer Vergrößerung des Abstandes zum nächsten Nachbarn auch immer eine Verkleinerung der größten Überlappungsfläche verbunden ist, betrachten wir die Änderung der Gesamtüberlappungsfläche bei Verschiebung des mittleren Kreises in Abbildung 37. Die Formel für die Überlappungsfläche A zweier gleich großer Kreise mit Radius r lautet:

$$A = 2 \left(r^2 \arccos \left(1 - \frac{h}{r} \right) - (r - h) \sqrt{2rh - h^2} \right)$$

Bei Verwendung der Bezeichnungen aus Abbildung 38 ergibt sich durch Einsetzen von $h_1 = r - 0.5d_1$ und für $h_2 = r - 0.5(d - d_1)$ für die Gesamtüberlappungsfläche in Abhängigkeit zum Abstand d_1 aus:

$$\begin{aligned} A_{ges} = & 2 \left(r^2 \arccos \left(1 - \frac{r - 0.5d_1}{r} \right) - 0.5d_1 \sqrt{2r(r - 0.5d_1) - (r - 0.5d_1)^2} \right) \\ & + 2 \left(r^2 \arccos \left(1 - \frac{r - 0.5(d - d_1)}{r} \right) \right. \\ & \left. - 0.5(d - d_1) \sqrt{2r(r - 0.5(d - d_1)) - (r - 0.5(d - d_1))^2} \right) \end{aligned}$$

Abbildung 38 zeigt die Kurvenschar mit dem Parameter d der Gleichung für Einheitskreise. Für d sind sieben Werte zwischen 0 und 3.5 dargestellt. Ab einem Abstand von $d=4$ bei $r=1$ ist es nicht mehr möglich, dass der mittlere Kreis sich mit beiden äußeren Kreisen überlappt. Der Definitionsbereich ist für $d = 2r$ am größten, da sich bei diesem Abstand die beiden äußeren Kreise berühren und sich der mittlere Kreis daher für jeden Wert von d_1 zwischen 0 und $2r$ mit beiden überschneidet. Für kleinere d ist der Definitionsbereich durch d nach oben begrenzt und für größere d muss $d_1 > d - 2r$ gelten, damit sich der mittlere Kreis mit beiden äußeren Kreisen überlappt.

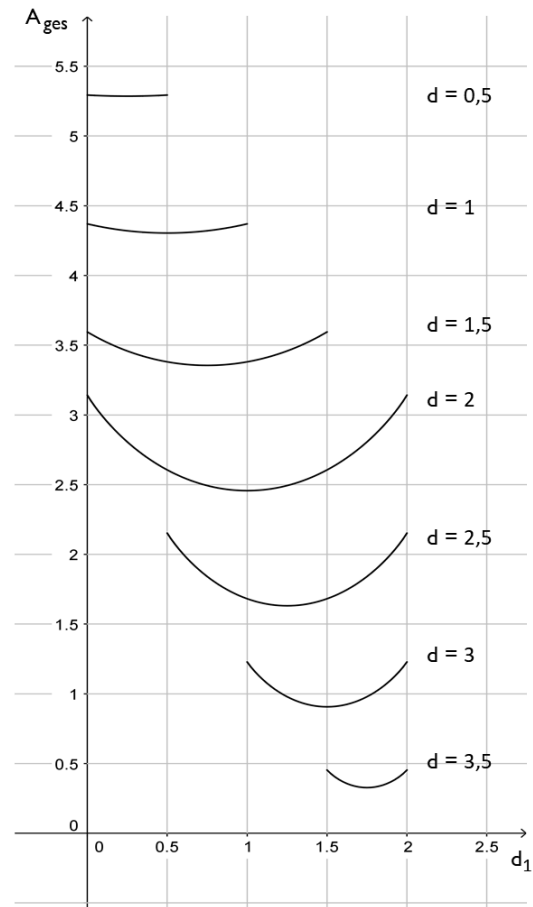


Abb. 38: Größe der Überlappungsfläche in Abhängigkeit des Abstands

Schwerpunkt-Heuristik

Die Idee für die zweite Heuristik stützt sich auf Schwerpunkt-Voronoi-Diagramme (SVD) (siehe Abschnitt 2.2). Wie im vorigen Abschnitt gezeigt, sind auch die Überlappungsflächen minimal, wenn die Abstände zwischen überlappenden Kreisen gleich groß sind. Daher ist bei einer Anordnung der Punkte, bei der die Energie minimal ist, auch die Überlappungsfläche der Signaturen minimal.

Für die Iterationsschritte besteht jedoch kein direkter Zusammenhang zwischen der Energie und der Überlappungsfläche. In Abbildung 39 ist ein Beispiel einer Anordnung gezeigt, bei der durch die Verschiebung der Kreise auf die Schwerpunkte ihrer Voronoi-Zellen die Überlappungsfläche zunimmt, obwohl die Energie geringer ist. Auf der linken Seite ist die Anordnung vor der Verschiebung abgebildet sowie die Schwerpunkte der Voronoi-Zellen in grün. Auf der rechten Seite ist die Situation nach der Verschiebung der Mittelpunkte auf den Schwerpunkt dargestellt.

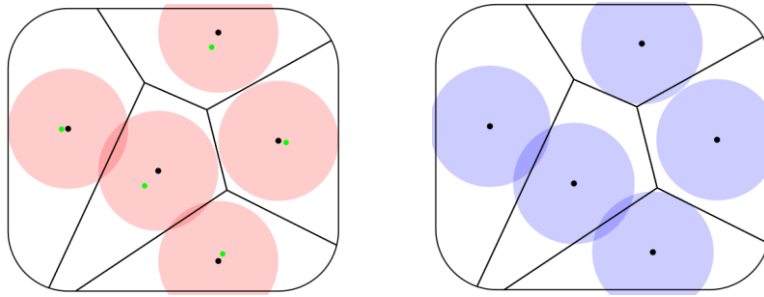


Abb. 39: Vergrößerung der Überlappungsfläche durch Verschiebung auf den Schwerpunkt

Folglich ist nicht garantiert, dass die Überlappungsfläche in jedem Iterationsschritt verringert wird, jedoch ermöglicht dies, dass lokale Minima überwunden werden, die bei Anwendung der GIK-Heuristik nicht aufgelöst werden können. Ein Beispiel für einen solchen Fall ist in Abbildung 40 zu sehen. Auf der linken Seite ist die Anfangsanordnung der Punkte sowie der GIK in türkis und der größte innere Kreis mit dem Schwerpunkt als Mittelpunkt in grün dargestellt. Im mittleren Bild ist das Ergebnis bei Verwendung der GIK-Heuristik und im rechten das Ergebnis der Schwerpunkt-Heuristik nach je 100 Iterationen zu sehen. Sowohl die Gesamtüberlappungsfläche als auch die größte Überlappungsfläche ist im rechten Bild geringer.

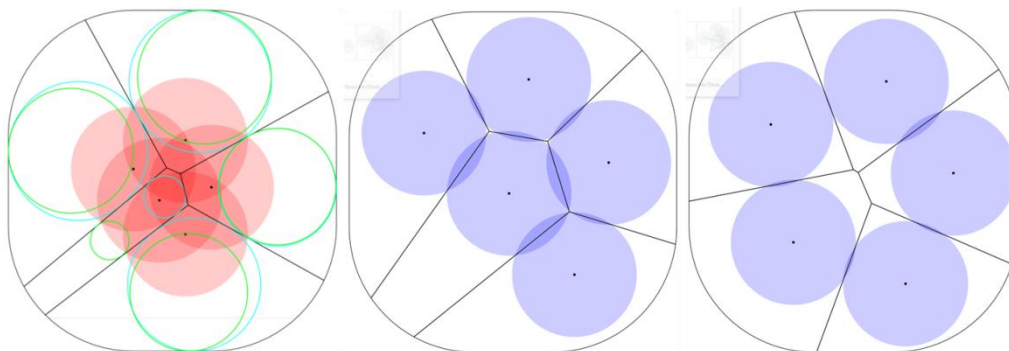


Abb. 40: Überwindung eines lokalen Minimums durch die Schwerpunkt-Heuristik

7.1.2 Beschränkung der Fläche der Voronoizellen

Bei länglichen Voronoizellen, wie in Abbildung 41 im linken Bild zu sehen, führen beide Heuristiken nicht zu einer überlappungsfreien Lösung, obwohl die Überlappung aufgelöst werden könnte. Der Grund hierfür ist, dass der Mittelpunkt des größten Kreises bzw. der Schwerpunkt weiter von der Position des POIs entfernt ist als der Schwellwert und somit die nach der Heuristik beste Position nicht erreicht werden kann.

Ziel ist es daher, die Menge der durch die Heuristiken berechneten Positionen einzuschränken, ohne potentiell optimale Positionen auszuschließen. Um dieses zu erreichen, wird um jeden POI ein achsenparalleles Quadrat konstruiert, dessen Kantenlänge gleich der vierfachen maximalen Verschiebungsdistanz ist und dessen Mittelpunkt auf den POI liegt. Anschließend wird für jeden POI das Schnittpolygon der Voronoizelle und des Quadrates berechnet und an Stelle der Voronoizelle für die Berechnung des Verschiebungsvektors verwendet. Dies ist in Abbildung 41 auf der rechten Seite für die GIK-Heuristik dargestellt. Die Komplexität für die Berechnung des Schnittes zweier konvexer Polygone ist $O(n)$ [OCO+82] und das Schnittpolygon ist stets konvex, damit bleibt die Komplexität für die Berechnung des Verschiebungsvektors insgesamt unverändert.

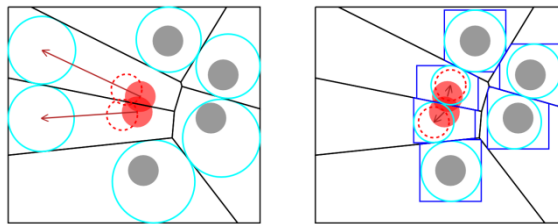


Abb. 41: Einfluss der Begrenzung der Voronoizelle auf die Verschiebungsvektoren

Die Richtung des Verschiebungsvektors ändert sich abhängig vom Größenverhältnis zwischen Voronoizelle und Quadrat, jedoch bleiben die in Abschnitt 7.1.1 getroffenen Aussagen über den GIK-Verschiebungsvektor gültig. Der Grund hierfür ist, dass die Voronoizelle durch das Quadrat nur dort verkleinert wird, wo der Abstand zwischen der Voronoikante und dem POI kleiner als der zweifache Radius der Signatur ist. Folglich bleibt zum einem zwischen einem POI und seinem nächsten Nachbarn immer die Voronoikante erhalten, falls eine Verschiebung durchzuführen ist und zum anderen ist der Abstand zwischen dem POI und den Nachbarn zwischen denen eine Kante des Quadrates liegt auch nach der Verschiebung mindestens so groß wie der Radius.

Die Berücksichtigung von Hintergrundobjekten bei der Verdrängung kann ebenfalls durch eine Beschränkung der Voronoizellen realisiert werden. Allerdings kann in den meisten Fällen nicht das Hintergrundobjekt selbst mit der Voronoizelle verschnitten werden, da das resultierende Polygon konvex sein muss, damit das in dieser Arbeit entwickelte Verdrängungsverfahren durchgeführt werden kann. Bei linienförmigen Objekten kann, durch die Verwendung einer Geraden als Approximation, die Konvexität des Schnittpolygons gewährleistet werden. Durch die Wahl der Approximation lassen sich Prioritäten in Bezug auf die Überdeckung von POI-Signaturen und Hintergrundobjekten treffen, wie in Abbildung 42 dargestellt ist. In der oberen Reihe ist jeweils die Situation vor der Verdrängung abgebildet. Die Zelle um die POIs ist zusätzlich durch die violetten Linien beschränkt und der GIK ist in türkis eingezeichnet. Jeweils darunter ist das Ergebnis der Verdrängung zu sehen, wobei die Signaturen, die weder eine Signatur noch die Hintergrundobjekte über-

decken grün eingefärbt sind. Im linken Bild werden die Autobahn und der Fluss nicht bei der Berechnung des Verschiebungsvektors berücksichtigt, so dass eine Überdeckung des Hintergrundes möglich wenn auch nicht notwendig ist. So würde nach der Verdrängung die obere Signatur den Fluss überdecken. Liegt ein Teil des Hintergrundobjekts auf der Approximation und der Rest weiter vom POI entfernt als die Approximation, so werden diese gleich gewichtet (Abbildung 42, mittleres Bild), in dem Sinne, dass die Hintergrundobjekte nur dann überdeckt werden, wenn die Überlappung der Signaturen nicht aufgelöst werden kann. Außerdem ist garantiert, dass der Mittelpunkt der Signatur nicht auf die andere Seite des Objekts verschoben werden kann. Liegt ein Teil des Objekts näher am POI, so wird eine Überdeckung des Objektes teilweise ermöglicht. Soll sichergestellt werden, dass das Hintergrundobjekt nicht überdeckt wird, so ist der minimale Abstand zwischen Approximation und Objekt auf die maximale Verschiebungsdistanz zu setzen (Abbildung 42, rechtes Bild). Hierbei werden die Signaturen weiter verschoben, als zur Auflösung der Überlappungen notwendig ist und sofern möglich wird der Abstand der Signatur zum Hintergrundobjekt auf die maximale Verschiebungsdistanz erhöht. Falls dies nicht möglich ist, werden auch die Überlappungen der Signaturen nicht aufgelöst.

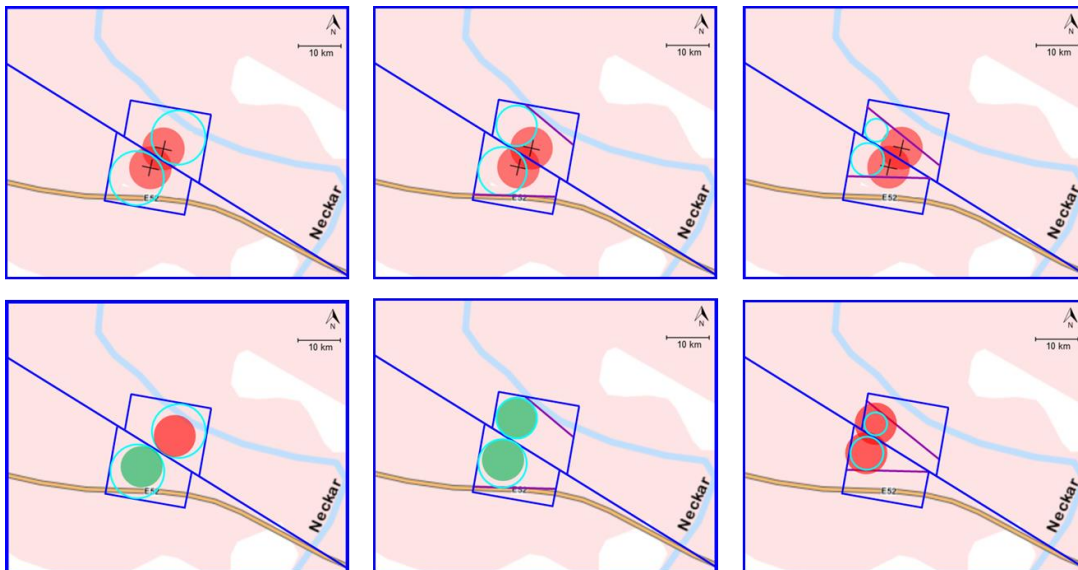


Abb. 42: Berücksichtigung von Hintergrundobjekten bei der Verdrängung

7.1.3 Zerlegung in unabhängige Teilmengen

Da die Distanz zwischen einem verschobenen Symbol und den Koordinaten des POIs streng beschränkt ist, ist es möglich Fixpunkte zu bestimmen, d.h. Punkte die keinesfalls verdrängt werden müssen. Des Weiteren ist es möglich die Punktmenge in disjunkte Untermengen zu zerlegen, die unabhängig voneinander bearbeitet werden können. Für beide Schritte kann ein Voronoi-Diagramm als Hilfsstruktur

verwendet werden. Um die Fixpunkte zu bestimmen, wird für jeden POI die Distanz zu seinem nächsten Nachbarn bestimmt. Falls diese Distanz größer ist, als der doppelte Durchmesser der Signatur, ist es nicht möglich, dass es durch Verschiebung der POIs zur Überlappung mit dem nächsten Nachbarn (und somit auch mit keinem anderen POI) kommt.

Um die Punktmenge in disjunkte Untermengen zu zerlegen, wird zunächst die Menge der Voronoikanten bestimmt, die für die Zerlegung verwendet werden können. Hierzu wird für jede Kante die Distanz zu den beiden Orten berechnet, durch die sie definiert ist. Falls die Distanz kleiner als der Durchmesser der Signatur ist, ist es möglich, dass sich die Symbole nach der Verschiebung überlappen. Folglich verbleiben beide Orte in einer Untermenge und die Kante darf nicht zur Aufteilung des Diagramms verwendet werden und wird deswegen gelöscht. In Abbildung 43 auf der linken Seite sind die verbliebenen Kanten eines Voronoi-Diagramms von 513 Orten (Flughäfen in Europa) dargestellt. Anschließend wird das Diagramm rekursiv entlang eines willkürlich gewählten Pfades, der von einem Grenzpunkt zu einem anderen führt, aufgeteilt, siehe Algorithmus 2. Als Eingabe erhält der Algorithmus einen beliebigen Grenzpunkt mit mindestens einer adjazenten Kante und den Graph der verbliebenen Kanten. In der ersten Iteration sind nur die Schnittpunkte des Voronoi-Diagramms mit dem Clip-Polygon Grenzpunkte. In jeder weiteren Iteration werden jeweils die Stützpunkte des gefundenen Pfades zu der Menge der Grenzpunkte hinzugefügt und die Kanten des Pfades aus dem Graph der verbliebenen Kanten gelöscht. Das Ergebnis der Zerlegung ist auf der rechten Seite der Abbildung 43 gezeigt: Die Voronoizellen der POIs jeder Untermenge sind unterschiedlich eingefärbt und die Voronoizellen der Fixpunkte sind weiß.

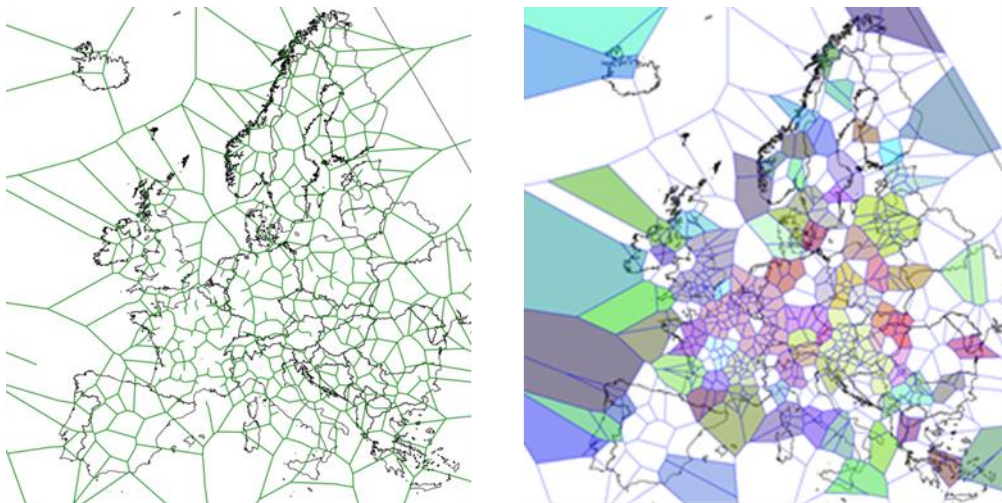


Abb. 43: Aufteilung der Punkte in Europa in unabhängige Teilmengen

```
procedure findPath(Node p, Graph g)
2:  branch  $\leftarrow$  Stack of nodes
3:  splitPath  $\leftarrow$  List of nodes
4:  finished  $\leftarrow$  false
5:  splitPath.add(p)
6:  while not finished do
7:      if p has more than one adjacent node do
8:          branch.push(p)
9:      end if
10:     np  $\leftarrow$  adjacent node of p
11:     remove edge between p and np from g
12:     if path contains np or np is border point do
13:         finished  $\leftarrow$  true
14:     end if
15:     path.add(np)
16:     if np has no adjacent node do
17:         p  $\leftarrow$  branch.pop()
18:         remove points from splitPath, that have been inserted after p
19:     else
20:         p  $\leftarrow$  np
21:     end if
22: end while
23: return path
```

Algorithmus 2: Berechnung eines Pfades zur Aufteilung der Punktmenge

Im Allgemeinen ist die Komplexität um einen Pfad von einem Grenzpunkt zu einem anderen zu finden linear zur Anzahl der Kanten. In diesem Fall ist das Auffinden aller Untermengen ebenfalls linear zur Anzahl der Kanten. Der Grund hierfür ist, dass jede Kante nur genau einmal betrachtet werden muss und dann gelöscht werden kann (Zeile 11), wie die folgende Untersuchung der einzigen drei möglichen Fälle, die in Abbildung 44 gezeigt sind, belegt:

- a) Die Suche endet an einem Grenzpunkt (Zeile 12).
- b) Der Pfad enthält einen Zyklus und der Algorithmus hält an sobald der Zyklus geschlossen ist (Zeile 12).
- c) Der Pfad endet nicht an einem Grenzpunkt, wegen gelöschter Kanten (Zeile 16).

In Fall a) werden alle betrachteten Kanten zur Aufteilung des Diagramms verwendet, d.h. nur ihre Endpunkte werden der Menge der Grenzpunkte hinzugefügt und die Kanten werden gelöscht. Gleiches gilt für Fall b). Der einzige Unterschied ist, dass das Diagramm in einen Teil innerhalb und einen Teil außerhalb des Kreises aufgeteilt wird anstelle eines linken und eines rechten Teils. In Fall c) werden alle Kanten nach der letzten Verzweigung aus dem Pfad gelöscht (Zeile 18) und die Suche wird von der letzten Verzweigung aus fortgesetzt und anschließend implizit erneut als a), b) oder c) klassifiziert. Werden zu jeder Kante die POIs durch die sie erzeugt werden und zu jedem POI die Kanten seiner Voronoizelle abgespeichert, so kann die Zuordnung von POIs zu den Gebieten in linearer Zeit erfolgen, sofern die Daten beim Entfernen einer Kante, die nicht zur Aufteilung verwendet werden darf, aktualisiert werden.

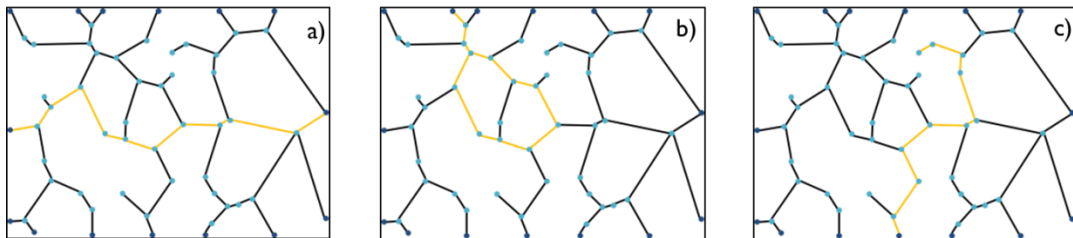


Abb. 44 Fallunterscheidungen bei der Suche nach Pfaden zur Zerlegung der Punktmenge

7.1.4 Auswahl der Parameter

Für das Verfahren sind zwei Parameter festzulegen: die zulässige maximale Verschiebungsdistanz und die Größe der Signatur. Prinzipiell können beide Parameter entsprechend den persönlichen Anforderungen/Präferenzen des Benutzers gewählt werden. Im Folgenden werden Ansätze für das Festlegen von Standardwerten bzw. Vorschläge für Parameteränderungen beschrieben, die zu einer Visualisierung mit weniger Überlappungen führen.

Eine Möglichkeit um eine in der Darstellung ablesbare Genauigkeit der Positionierung zu garantieren ist zu fordern, dass sich die Koordinaten jedes POIs innerhalb seiner Signatur befinden. Folglich beträgt die Distanz zwischen der tatsächlichen Position und des Mittelpunkts der Signatur höchstens den Radius der Signatur. Zur Illustration dieser Bedingung ist in Abbildung 45 der Stuttgarter Flughafen durch eine grüne Punktsignatur symbolisiert. Das Kreuz gibt die exakte Position an und der gestrichelte Kreis markiert das Gebiet, in dem die Signatur verschoben werden kann unter der Bedingung, dass die Signatur die Koordinaten des Flughafens enthält. Im linken Bild liegt der Mittelpunkt der Signatur auf der exakten Position des Flughafens, im mittleren Bild ist die Signatur ein wenig verschoben und im rechten Bild ist sie um die maximal zulässige Distanz verschoben.



Abb. 45: Begrenzung der Verschiebungsdistanz auf den Signaturreadius

Der Standardwert für die Signaturgröße kann aus den technischen Rahmenbedingungen abgeschätzt werden, denn der anzuzeigende Kartenausschnitt ist durch die Daten vorgegeben und der Maßstab durch das Display des verwendeten Gerätes beschränkt. Die Größe der Signatur ist durch die graphische Mindestgröße festgelegt, die sich bei einfachen Signaturen auf die Größe der Signatur und bei komplexen zusammengesetzten Signaturen auf die Lesbarkeit der einzelnen Komponenten bezieht.

Mit Hilfe des Voronoi-Diagramms ist es auch möglich in linearer Zeit zu bestimmen wie groß der Radius der Signaturen sein darf, so dass eine überlappungsfreie Darstellung ohne Verschiebung möglich ist. Hierzu wird der Punkt bestimmt der den geringsten Abstand zu einer seiner Voronoikanten hat und dieser Abstand wird dann als Radius gewählt. Wenn nur eine Iteration des Verfahrens ausgeführt werden soll und die Darstellung überlappungsfrei, so kann für jede Voronoizelle der GIK bestimmt werden und der Signaturreadius gleich dem Radius des kleinsten GIK gesetzt. Falls die Verschiebungsdistanz beschränkt sein soll, sind zuvor die Zellen, wie in Abschnitt 7.1.2 beschrieben zu begrenzen.

Falls das Ergebnis des Verdrängungsverfahrens nicht zufriedenstellend ist, also zu viele Überlappungen nicht aufgelöst werden können, ist es möglich auf Grundlage des Ergebnisses die maximale Größe des Signaturreadius abzuschätzen, für den eine überlappungsfreie Darstellung wahrscheinlich möglich ist. Hierzu wird die Größe der Fläche der Signatur, die am wenigsten sichtbar ist, sowie der Radius der zugehörigen Kreisfläche berechnet. Da alle Signaturen verkleinert werden, wird jede Signatur um die Hälfte der Differenz des neu berechneten zum ursprünglichen Radius verkleinert. Anschließend wird das Verfahren erneut mit den veränderten Parametern durchgeführt.

7.2. Verallgemeinerung des Verfahrens

Das in diesem Kapitel vorgestellte Verdrängungsverfahren ist speziell auf gleichgroße kreisförmige Signaturen zugeschnitten. Einige der in dem Abschnitt 7.1.1 nachgewiesenen Eigenschaften der Heuristiken, gelten nicht für andere Signaturformen, bzw. auch dann nicht, wenn eine konservative oder progressive kreisförmige Approximation verwendet wird. Bei der Verwendung von progressiven Approximationen kann bei einer überlappungsfreien Anordnung der Approximationen nicht garantiert werden, dass sich die Signaturen nicht überlappen. Bei konservativen Approximationen werden hingegen Signaturen weiter verschoben als notwendig, da sich die Approximationen überlappen können, ohne dass sich die Signaturen überlappen. Des Weiteren kann sich auch bei der Verwendung der GIK-Heuristik die größte Überlappungsfläche vergrößern, wie in Abbildung 46 zu sehen ist. Die Gesamtüberlappungsfläche der Approximationen ist im rechten Bild kleiner als im linken, die rautenförmigen Signaturen überlappen sich jedoch auf der rechten Seite stärker.

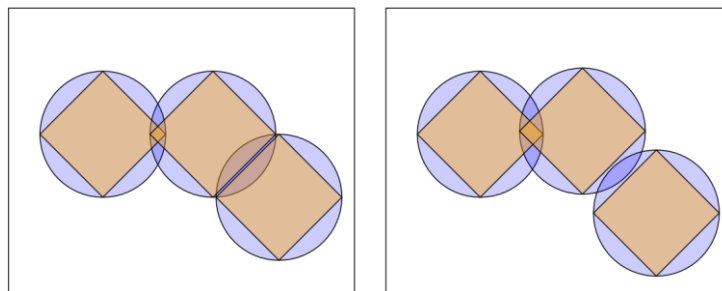


Abb. 46: Vergrößerung der Überlappungsfläche von Signaturen durch Verkleinerung der Überlappungsfläche ihrer konservativen Approximationen

Um auch auf andere Signaturformen und/oder unterschiedliche Signaturgrößen angewendet zu werden, ist eine Anpassung des Verfahrens notwendig. Hierzu können drei grundlegend verschiedene Ansätze verfolgt werden. Zum einem kann ein Vorverarbeitungsschritt durchgeführt werden, um die gegebenen Signaturen auf Kreise abzubilden, bevor das Verfahren unverändert angewendet wird und im Anschluss das Ergebnis zurücktransformiert wird. Als zweite Möglichkeit kann an Stelle des Voronoi-Diagramms für Punkte eine andere Variante verwendet werden, sofern diese Variante die für das Verfahren relevanten Eigenschaften des Voronoi-Diagramms aufweist. Wenn diese beiden Möglichkeiten nicht zur Verfügung stehen, ist als dritter Ansatz eine Veränderung der Voronoi- bzw. Powerzellen denkbar, die durchgeführt wird, bevor jeweils der Punkt berechnet wird, in deren Richtung der POI verschoben wird. Ein Nachteil dieser Variante ist, dass die Form der Signatur nicht bei der Partitionierung des Raumes berücksichtigt wird, sondern nur für die Platzierung innerhalb seiner Zelle. Im Folgenden werden die drei Ansätze an je einem konkreten Beispiel skizziert.

7.2.1. Symbole unterschiedlicher Größe

Wie bereits erwähnt, ist die grundlegende Idee des in dieser Arbeit entwickelten Verfahrens, jedem Punkt eine eindeutige Fläche zuzuordnen, die groß genug ist, um die Signatur komplett zu enthalten. Werden unterschiedlich große Signaturen verwendet, so ist der unterschiedliche Platzbedarf schon bei der Erstellung der Parquettierung zu berücksichtigen, da ansonsten beide Heuristiken zu keinem guten Ergebnis führen, wie in Abbildung 47 zu sehen ist. Die Signatur in der Mitte des Bildes auf der linken Seite, würde sowohl nach der GIK-Heuristik (türkiser Punkt) als auch nach der Schwerpunkt-Heuristik (hellgrüner Punkt) in Richtung der beiden größeren Signaturen verschoben werden, obwohl eine Auflösung der Überlappung durch Verschiebung der mittleren Signatur in Richtung der kleineren Signaturen ohne Verschiebung der anderen Signaturen möglich ist, wie im mittleren Bild dargestellt. Der Grund hierfür ist, dass die implizite Annahme, dass auch der freie Raum zwischen zwei benachbarten Signaturen durch die Voronoikante gleichmäßig zwischen ihnen aufgeteilt wird, nicht mehr gilt, wie beim Vergleich des rechten Bildes mit dem linken Bild zu sehen ist.

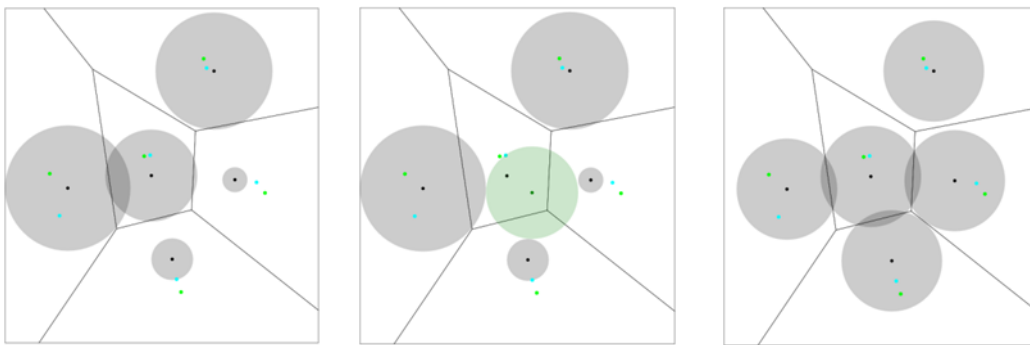


Abb. 47: Berechnung der Verschiebungsvektoren für unterschiedlich große Signaturen mittels des Voronoi-Diagramms

Eine Verallgemeinerung des Voronoi-Diagramms, bei dem es sich bei den Orten nicht um Punkte sondern um Sphären handelt, wird Power-Diagramm genannt [Au87]. Bei einem Power-Diagramm wird jeder Punkt der Ebene dem Ort zugeordnet, für den der sogenannte Power-Wert $v = d^2 - r^2$ minimal ist. Hierbei ist d die euklidische Distanz zwischen dem Punkt und dem Mittelpunkt und r der Radius der Sphäre. Wie beim Voronoi-Diagramm handelt es sich bei den Zellen um konvexe Polygone. Allerdings gilt der Nachweis nicht, dass durch die GIK-Heuristik die größte Überlappungsfläche in jedem Fall verkleinert wird, da der Zuwachs der Überlappungsfläche nicht nur durch die Abstände der Mittelpunkte, sondern auch von den beteiligten Radien abhängt. So wird die Gesamtüberlappungsfläche unter Umständen vergrößert wenn ein Kreis in Richtung des Kreises mit größeren Abstand verschoben wird, falls der Radius dieses Kreises deutlich größer ist, als der Radius des Kreises zu dem der Abstand vergrößert wird. Jedoch existieren für

Schwerpunkt-Power-Diagramme vergleichbare Ergebnisse wie für Schwerpunkt-Voronoi-Diagramme, so dass die Aussagen zur Schwerpunkt-Heuristik gleich bleiben [BR15]. In Abbildung 48 ist im linken Bild das Voronoi-Diagramm der Mittelpunkte der Kreise in schwarz und das Power-Diagramm der Kreise in blau eingezeichnet. Im mittleren Bild ist das Ergebnis des Verfahrens bei Verwendung des Power-Diagramms dargestellt und rechts bei Verwendung des Voronoi-Diagramms.

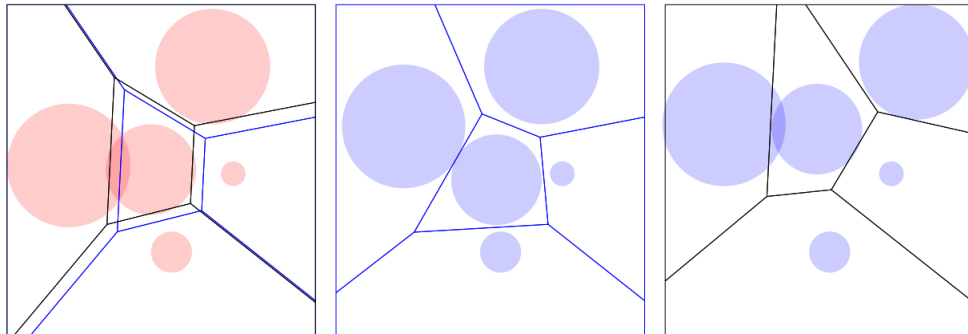


Abb. 48: Vergleich der Verdrängung unterschiedlich großer Signaturen mittels Power-Diagramm (mitte) und Voronoi-Diagramm (links)

7.2.2. Lokale Änderung der Zellgeometrie zur Berechnung der Verschiebungsrichtung

Die Idee bei dieser Variante ist es, die Zellgeometrie vor der Berechnung des Verschiebungsvektors zu verändern, um zu vermeiden dass Signaturen durch das Verfahren verdrängt werden, obwohl sich nur ihre konservativen Approximationen überlappen. Ziel ist es also die Platzierung der Signatur innerhalb ihrer Voronoi- bzw. Powerzelle zu verbessern. Als Beispiel wird die aus zwei Halbkreisen zusammengesetzte Signatur aus Abschnitt 6.4 betrachtet.

Die kreisförmige konservative Approximation der Signatur entspricht dem Vollkreis des größeren Halbkreises. Zur Veranschaulichung des Vorgehens ist in Abbildung 49 a) die Platzierung der Signatur innerhalb ihrer in blau gezeichneten Zelle dargestellt, wenn man die Position mittels der Approximation berechnet. Die Signatur ragt dabei aus der Zelle heraus, obwohl eine Platzierung innerhalb der Zelle möglich wäre (siehe Abbildung 49 b)). Ziel ist es nun die Zelle so zu vergrößern, dass aus einer Platzierung, bei der die Approximation komplett innerhalb der vergrößerten Zelle liegt, folgt, dass auch die Signatur komplett innerhalb der Originalzelle liegt. Für das Beispiel aus Abbildung 49 kann dies erreicht werden, indem die untere Kante um die Differenz der beiden Radien nach unten verschoben wird. In Abbildung 49 c) ist die optimale Platzierung der Approximation innerhalb der vergrößerten Zelle und in d) die tiefste Position bei der die Approximation noch komplett innerhalb der vergrößerten Zelle liegt dargestellt.

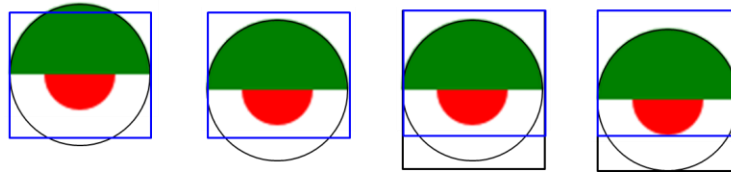


Abb. 49: Positionierungen der Halbkreis-Signatur innerhalb einer rechteckigen Zelle

Weiterhin muss die vergrößerte Zelle ebenfalls konvex sein, damit die Berechnungen der Verschiebungsvektoren erfolgen kann, wie in Abschnitt 7.1.1. beschrieben. Um dies für allgemeine konvexe Zellen umzusetzen, werden die Zellen zunächst an den Stützpunkten die am weitesten links und am weitesten rechts liegen in zwei Linienzüge, einen oberen und einen unteren aufgeteilt. Wenn der obere Halbkreis kleiner als der untere Halbkreis ist, werden die Stützpunkte des oberen Linienzuges senkrecht nach oben verschoben, andernfalls die des unteren Linienzuges senkrecht nach unten. Start- und Endpunkt des Linienzuges bleiben unverändert. Die Distanz um die ein Stützpunkt verschoben werden darf, hängt von der Steigung der Kante ab. Ist die Steigung geringer als die Steigung der Tangente g an den kleineren Halbkreis durch den Eckpunkt S des größeren Halbkreises (siehe Abbildung 50 links), so kann die Kante parallel um die Differenz der Radien verschoben werden. Wenn die Steigung größer ist, könnte die Ecke des unteren Halbkreises außerhalb der Originalzellgeometrie liegen, obwohl die Approximation komplett in der vergrößerten Zelle liegt. Aus diesem Grund muss die parallele Verschiebungsdistanz um die Entfernung d des Eckpunktes S von der Tangente g an den kleineren Halbkreis verringert werden (siehe Abbildung 50 rechts).

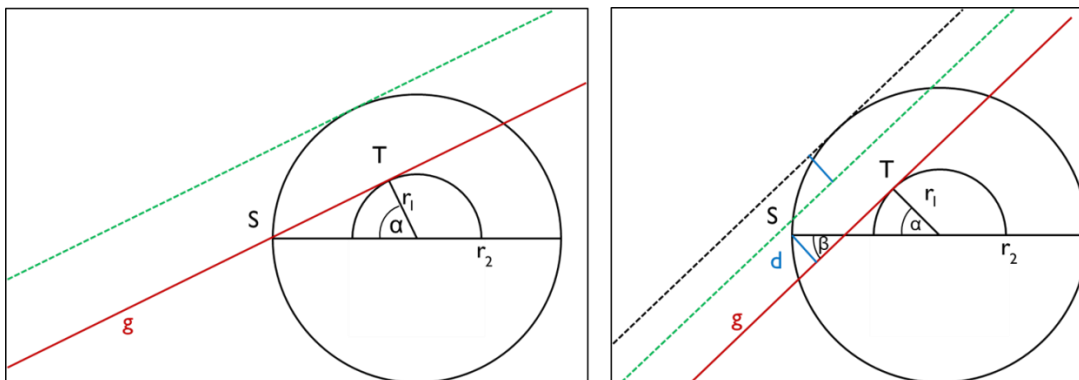


Abb. 50: Berechnung der Distanz zur Verschiebung der Kanten der Zellgeometrie

Anschließend wird für jeden der Stützpunkte der vertikale Abstand zu beiden parallel verschobenen inzidenten Kanten berechnet und der Stützpunkt um den kleineren der beiden Abstände senkrecht nach oben verschoben. Die verschobenen Stützpunkte bilden die neue Zellgeometrie.

7.2.3. Verdrängung elliptischer Symbole

Um eine Ellipse auf einen Kreis abzubilden, ist eine Stauchung entlang der längeren Halbachse, bzw. Streckung entlang der kürzeren Halbachse durchzuführen. Diese Abbildung lässt sich durch eine Rotation um den Ursprung und eine Parallelstreckung realisieren, d.h. es handelt sich um eine affine Abbildung, die sich durch folgende Matrix darstellen lässt:

$$R = \begin{pmatrix} \frac{a}{b} \cos \alpha & \frac{a}{b} \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix},$$

wobei α der Winkel zwischen der x-Achse und der Hauptachse der Ellipse ist. Für die Umkehrfunktion, die den Kreis in die ursprüngliche Ellipse zurücktransformiert, wird die Inverse der Matrix R verwendet. Durch diese Transformationen wird die Topologie der Objekte nicht verändert, da beide Funktionen bijektiv und stetig sind und es sich daher bei der Abbildung um einen Homöomorphismus handelt. Sind mehrere identische Ellipsen in einer Ebene angeordnet, so wird durch eine passende Wahl von (a, b, φ) jede dieser Ellipsen in einen Kreis transformiert.

In Abbildung 51 ist das Vorgehen zum Verdrängen elliptischer Signaturen veranschaulicht. Auf der linken Seite 1) sind die in Geonames vorhandenen Burgen im Südwesten Großbritanniens, durch elliptische Symbole markiert. In 2) ist der Raum in Richtung der y-Achse gestaucht, in 3) ist das Ergebnis der Verdrängung im gestauchten Raum zu sehen und auf der rechten Seite 4) ist das Ergebnis nach der Rücktransformation abgebildet.

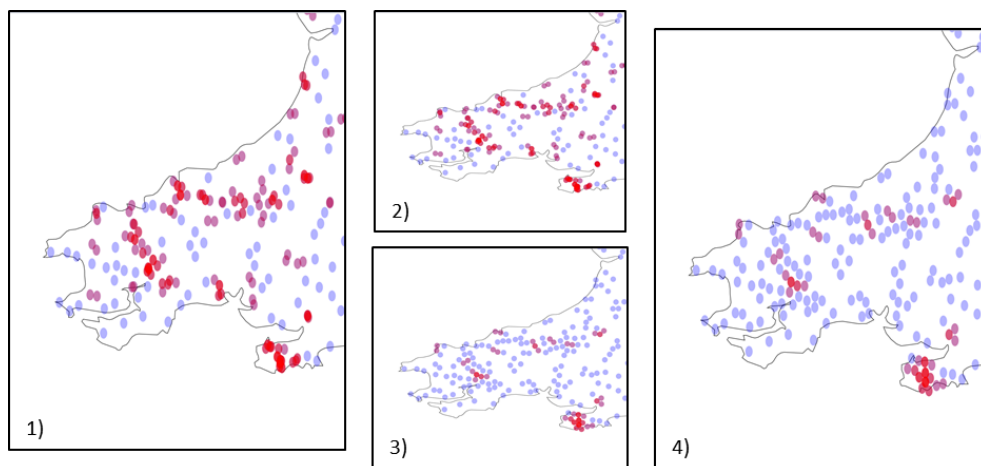


Abb. 51: Verdrängung elliptischer Symbole. 1) Eingabe. 2) Transformierter Raum. 3) Verdrängte POIs. 4) Rücktransformierter Raum

Resultate der Experimente

Im folgendem werden die Experimente und Evaluationen, die im Rahmen dieser Arbeit durchgeführt wurden, präsentiert und die Ergebnisse analysiert. Die Evaluation des Signaturentwurfs beschränkt sich auf die Darstellung von Sentiments durch Punktsignaturen und erfolgt mittels einer Online-Umfrage, deren Ergebnisse in Abschnitt 8.1 vorgestellt werden. Anschließend wird in Abschnitt 8.2 das Verdrängungsverfahren für gleich große Signaturen in Bezug auf die Zerlegung der Punktmenge (8.2.2), die Auflösung der Überlappungen (8.2.4) und die Erhaltung der räumlichen Lage evaluiert (8.2.5). Hierfür werden zunächst einige der POIs entfernt, damit im Testdatensatz weniger Überlappungs-Konstellationen die nicht aufgelöst werden können enthalten sind (8.2.3). Es folgt ein Vergleich der Ergebnisse der Verdrängung für ungleichgroße Signaturen mittels Voronoi-Diagramms mit den Ergebnissen mittels Power-Diagramm (8.3) und der Anpassung der Zellgeometrie vor der Berechnung des Verschiebungsvektors (8.4). Des Weiteren werden für die Testdatensätze auf Grundlage der Ergebnisse des Verdrängungsverfahrens geeignete Signaturradien berechnet und die Ergebnisse in Abschnitt 8.5 ausgewertet.

8.1. Benutzerstudie zur Lesbarkeit der Punktsignaturen

Mit Hilfe der Online-Umfrage, die mit Hilfe der Umfrage-Software Unipark erstellt und ausgewertet wurde, werden folgende Fragestellungen untersucht:

- Welche der in Abschnitt 6.4 entworfenen Punktsignaturen werden präferiert?
- Welche Signaturen können am zuverlässigsten nach ihrer Größe sortiert werden?
- Gibt es einen Zusammenhang zwischen Präferenz und Sortierung?

Nach den demografischen Fragen werden zunächst für jede der Punktsignaturen-paare aus Abschnitt 6.4 drei Kombinationen, wie in Tabelle 3 zu sehen angezeigt, aus denen der Teilnehmer eine auswählen soll, die er/sie am geeignetsten für die Darstellung von Sentiments hält. Für die Fragen in Bezug auf die Sortierung nach Größe der Signaturen werden jeweils 7 Signaturen mit unterschiedlichen, manuell gewählten positiven und negativen Sentimentwerten überlappungsfrei auf einer rechteckigen Fläche platziert. Auf diese Weise wurden sechs Anordnungen erstellt, die in Abbildung 52 dargestellt sind. Für jedes Punktsignaturenpaar wird zufällig eine der sechs verschiedenen Anordnungen ausgewählt und der/die Teilnehmer/in gebeten zunächst die dargestellten positiven Werte nach ihrer Größe aufsteigend zu sortieren, anschließend die negativen und abschließend anzugeben, ob an einem

Ort die positiven oder die negativen Sentimentwerte überwiegen. Die Reihenfolge in der die Signaturen abgefragt werden ist randomisiert.

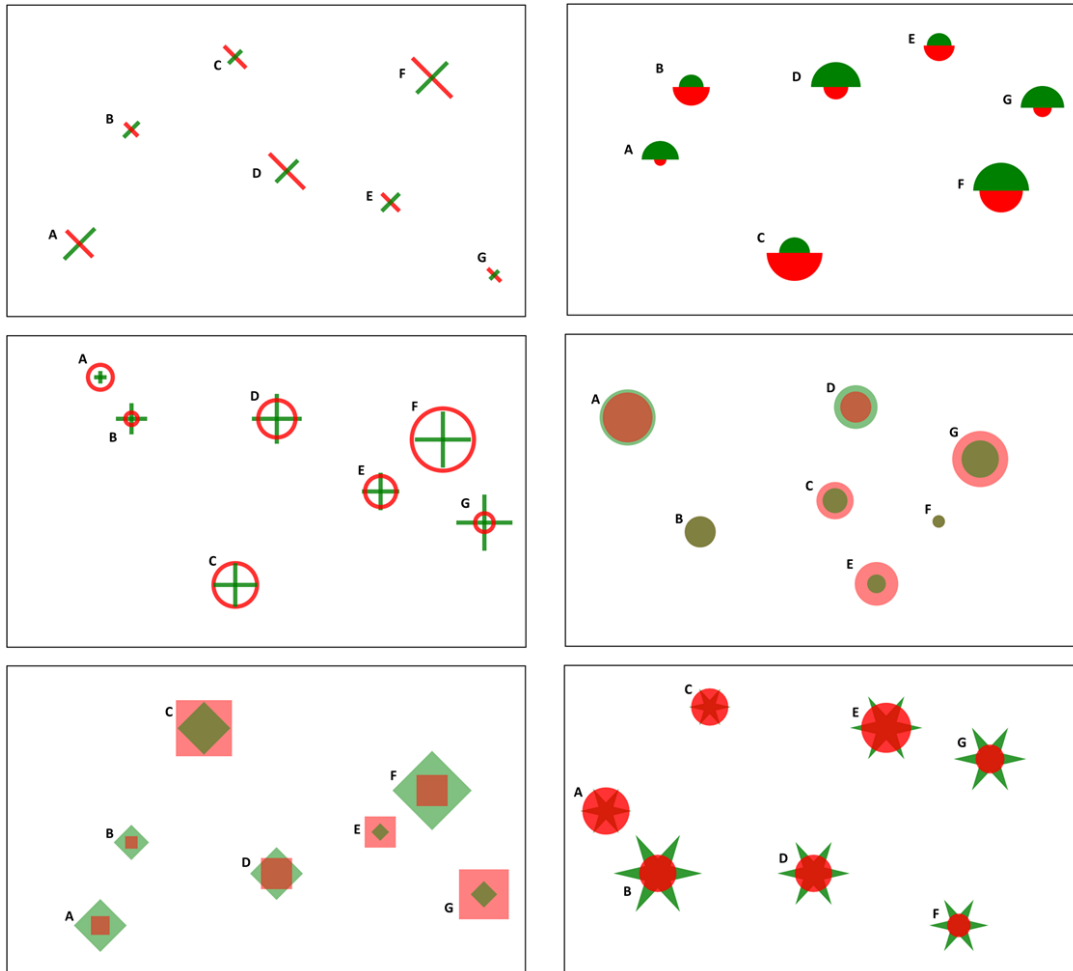








Abb. 52: Anordnungen und Signaturen für die Sortieraufgaben der Umfrage

Nach den Aufgaben zur Sortierung wird der Teilnehmer/in erneut gefragt welches der Signaturenpaare er/sie für die Darstellung von Sentiments auswählen würde. Insgesamt haben 46 Personen (21 Frauen und 25 Männer) an der Umfrage teilgenommen, von denen zwei bei der Auswertung nicht berücksichtigt wurden, da sie eine Rot-Grün-Schwäche haben. In Tabelle 3 ist angegeben wie viele Teilnehmer/innen, welches Signaturpaar präferieren und wie viele der Anordnungsaufgaben korrekt gelöst wurden. Allerdings liegen die Daten nach Signaturbeschriftung aggregiert vor, so dass eine Angabe von 50% entweder bedeuten kann, dass die Hälfte der Teilnehmer/innen die Aufgabe vollständig richtig sortiert hat, oder dass alle Teilnehmer/innen für die Hälfte der Symbole den korrekten Rang angegeben haben. Die Summe der präferierten Signaturenpaare ist zu Anfang und am Ende der Umfrage unterschiedlich und in beiden Fällen größer als die Teilnehmerzahl, da











































zehn der Teilnehmer/innen entgegen der Fragestellung mehr als ein Signaturenpaar ausgewählt haben.

Tabelle 3: Ergebnisse der Online-Umfrage

	Zu Anfang der Umfrage gewählt	Am Ende der Umfrage gewählt	Pot. Korrekte Angaben in%		
			positiv	negativ	Vergleich
	6	5	63,64	68,18	65,68
	5	2	81,40	88,37	88,0
	7	8	64,28	45,24	83,75
	9	10	58,14	88,37	88,59
	20	22	88,37	81,4	97,01
	4	0	79,07	88,37	94,02

Die Halbkreise wurden anfangs von 45,5% und am Ende der Hälfte der Teilnehmer/innen ausgewählt und sind damit, das am häufigsten gewählte Signaturen-paar. Die Werte für die Sortieraufgaben sind für die Halbkreise ebenfalls am höchsten. Allerdings wird das Signaturen-paar, das die zweithöchsten Werte aufweist, die Strichsignaturen, nach den Sortieraufgaben von niemandem gewählt. Insgesamt weisen die Ergebnisse darauf hin, dass es keinen Zusammenhang gibt, zwischen der Auswahl der Signaturen und dem Schweregrad, diese nach Größe zu ordnen. Des Weiteren fällt auf, dass bei gleichen Formen (Kreise, bzw. Striche) deutliche Unterschiede bei den Ergebnissen der Sortierung für positive und negative Werte vorliegen. Dies legt nahe, dass die Anordnung bzw. Kombination der Signaturen einen relevanten Einfluss auf die Größenwahrnehmung der Symbole hat. In Tabelle 4 ist für die 34 Teilnehmer/innen die sowohl zu Anfang (Zeile) als auch am Ende (Spalte) der Umfrage nur ein Signaturen-paar ausgewählt haben angegeben, ob bzw. wie sich die Präferenz geändert hat. Es sind nur 13 Personen bei ihrer anfänglichen Wahl geblieben, jedoch lässt sich auch in diesen Fällen kein klarer Zusammenhang zwischen der Richtigkeit der Sortiierungsergebnisse und der Präferenz erkennen.

Tabelle 4: Änderung der Präferenz

						
	  	  	  	  	  	  
  	1	0	1	2	0	0
  	1	0	0	1	0	0
  	0	0	2	1	3	0
  	0	0	2	2	3	0
  	2	1	1	2	8	0
  	0	0	0	0	1	0

8.2. Evaluation des Verdrängungsverfahrens

Eine prototypische Implementierung des in dieser Arbeit entwickelten Verfahrens erfolgt in Java mit Hilfe der Java Topology Suite (JTS), einer Open Source Java Bibliothek, die unter anderem Klassen zur Modellierung von Geometrien und der Berechnung von Voronoi-Diagrammen bereitstellt. Bei den Visualisierungen der Eingabe- und Ausgabedaten des Verdrängungsverfahrens handelt es sich um SVG-Dateien, die von dem in dieser Arbeit entwickelten Java-Programms erzeugt werden. Als Hintergrund werden Natural Earth Rasterdaten [KPI0] als Bild in die SVG-Datei eingebunden. Zunächst folgt eine Beschreibung der verwendeten Testdatensätze, an die sich eine Betrachtung der Zerlegung der Datensätze in unabhängig zu bearbeitende Gebiete anschließt. Danach werden die Ergebnisse der Experimente jeweils für beide Heuristiken vorgestellt und sowohl in Bezug auf die Lesbarkeit der Signaturen als auch die Erhaltung der räumlichen Anordnung evaluiert.

8.2.1. Testdatensätze

Für die Evaluation werden Daten des frei verfügbaren Gazetteers Geonames verwendet. In einem Vorverarbeitungsschritt werden Orte mit gleichem Objekttyp, gleichen Namen und geringem Abstand (< 5 km) eliminiert. Es werden fünf Objektarten für die Tests gewählt, die jeweils in ausreichender Anzahl und für die meisten Länder Europas vorliegen. Hierbei handelt es sich um Flughäfen (1.498), Hügel (162.177), Gipfel (25.921), Türme (5.094) und Städte (518.770) in Europa ohne Russland. In Abbildung 53 sind diese Orte jeweils durch eine 0,4 mm große Punktsignatur im Maßstab 1:20 Mio. abgebildet (nicht maßstäblich). Bei der Projek-

tion handelt es sich, wie auch bei allen folgenden kartenähnlichen Abbildungen in diesem Kapitel, um UTM 32N auf Basis des WGS-84-Ellipsoids.

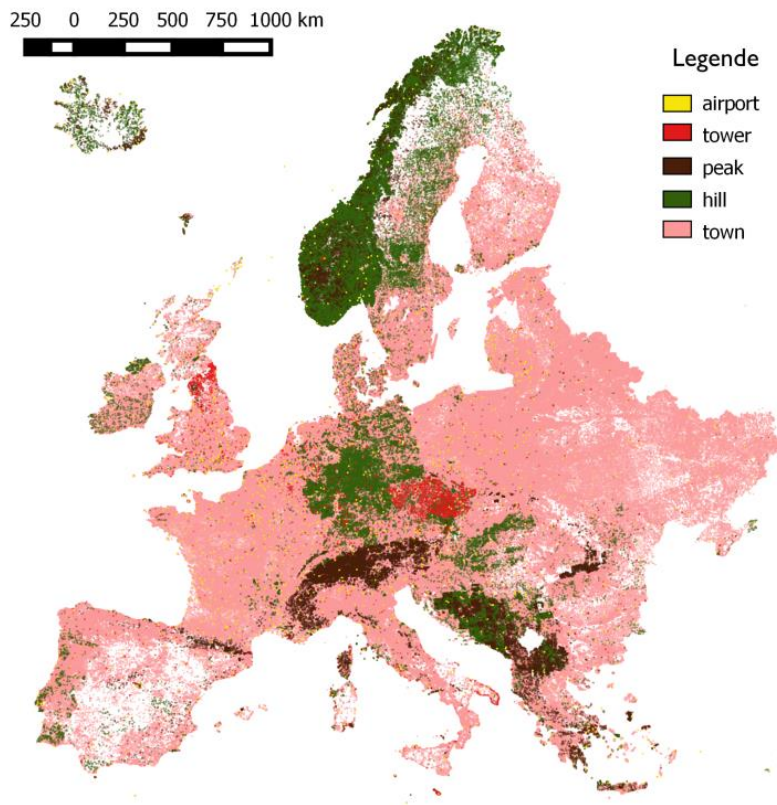


Abb. 53: Verteilung der POIs aus denen die Testdaten selektiert werden

Für jeden dieser Datensätze werden für zwei verschiedene Maßstäbe 1:5Mio. und 1:10Mio zufällig Objekte ausgewählt bis die Signaturfläche, bei einem Radius von 2mm, 10% der Gesamtfläche beträgt. Für den Maßstab 1:5 Mio ergeben diese Kriterien jeweils 2050 Objekte und für 1:10 Mio 513 Objekte. Auf diese Weise werden für jede Objektart und jeden Maßstab zehn Testdatensätze erzeugt, mit Ausnahme der Flughäfen im Maßstab 1:5Mio, da nicht genügend Flughäfen für mehrere Datensätze vorliegen. In Abbildung 54 sind zwei Datensätze, links Flughäfen und rechts Gipfel, für den kleineren Maßstab visualisiert.

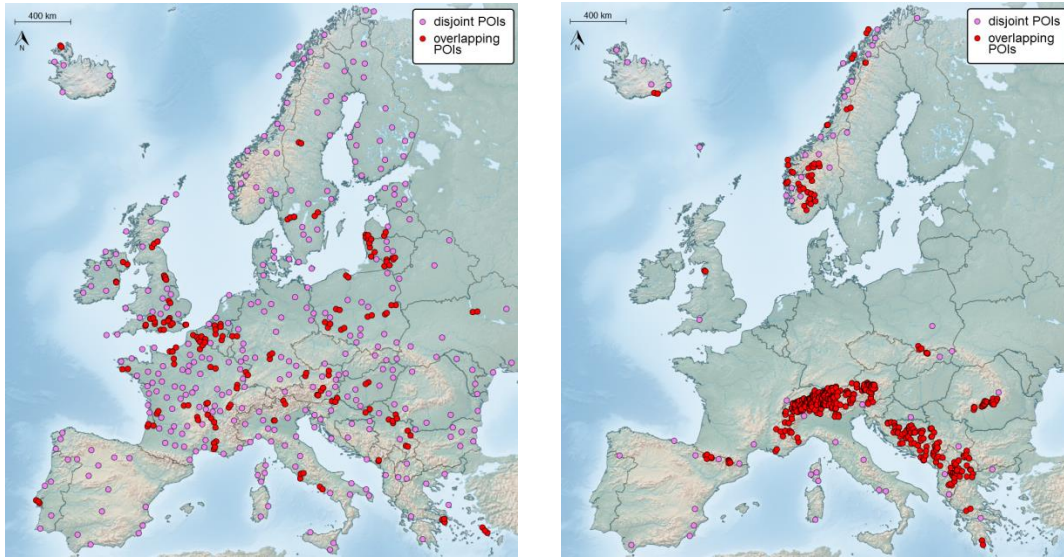


Abb. 54: Testdatensätze für Flughäfen (links) und Gipfel (rechts)

8.2.2. Zerlegung der Punktmenge

Die Anzahl und die Größe der Punktmenge, die sich jeweils unabhängig voneinander bearbeiten lassen (Cluster), variiert relativ stark zwischen den einzelnen Objekttypen, aber nur geringfügig zwischen den Datensätzen eines Objekttyps. In Abbildung 55 sind pro Objekttyp für fünf Datensätze des kleineren Maßstabs die Anzahl der Objekte, die sich in Clustern der jeweiligen Größe befinden dargestellt. Bei den Flughäfen- und den Städtedatensätzen enthalten alle Cluster weniger als 100 Objekte. Bei den Flughäfen ist mit über 20% Fixpunkten und mehr als der Hälfte aller Objekte in Clustern mit weniger als 10 Objekten der Anteil der kleinen Cluster etwas größer als bei den Städten. Bei den Gipfel- und Hügeldatensätzen liegt der Anteil an kleinen Clustern unter 20%, wobei bei den Hügeln größere Cluster (mit mehr als 300 Objekten) auftreten und bei den Gipfeln mehrere Cluster mit 50 bis 200 Objekten. Die Turm��atensätze enthalten um die 30% Fixpunkte und kleine Cluster, aber auch Cluster mit mehr als 200 Objekten.

In Bezug auf die Laufzeit ist ein Datensatz, der in mehrere kleinerer Punktmen- gen aufgeteilt werden kann, schneller zu berechnen, auch wenn die Größenord- nung gemäß O-Notation gleich bleibt. Denn es gilt:

$$\sum_{j=0}^m k_j \log \sum_{j=0}^m k_j > \sum_{j=0}^m (k_j \log k_j) \quad , \text{mit} \quad \sum_{j=0}^m k_j = n$$

$$\equiv \log \left(\left(\sum_{j=0}^m k_j \right)^{\sum_{j=0}^m k_j} \right) > \sum_{j=0}^m (\log(k_j^{k_j}))$$

$$\equiv \log \prod_{l=0}^m \left(\left(\sum_{j=0}^m k_j \right)^{k_l} \right) > \log \prod_{l=0}^m (k_l^{k_l})$$

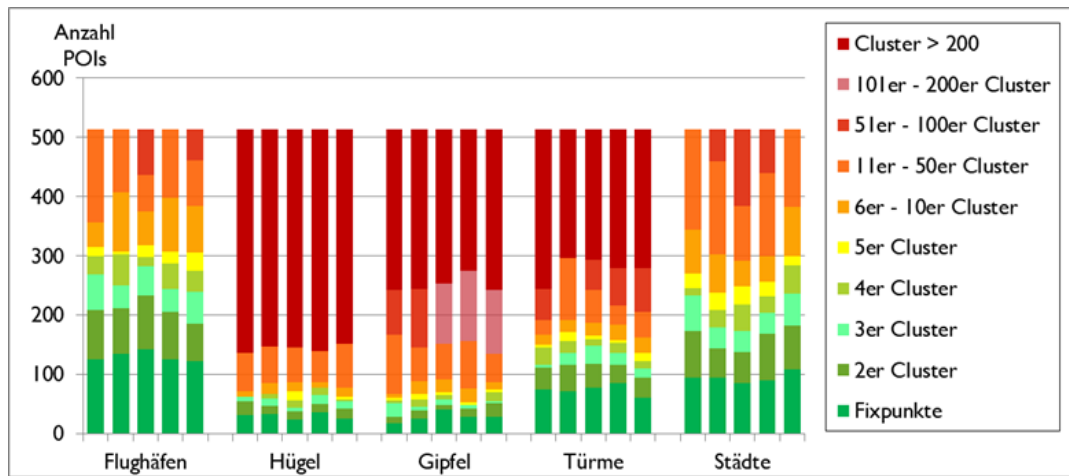


Abb. 55: Zerlegung der Datensätze in unabhängige Cluster

Des Weiteren reduziert sich die Anzahl der notwendigen Iterationen für den überwiegenden Teil der Cluster mit weniger als 10 Objekten auf unter 10, wie in Abbildung 56 zu sehen ist. Dort ist die Anzahl der Iterationen für die 10 Datensätze im kleineren Maßstab bei Anwendung der Schwerpunkt-Heuristik dargestellt. Fixpunkte werden bei den Verdrängungsverfahren nicht berücksichtigt (0 Iterationen) und für alle Cluster, die keine Überlappungen enthalten, endet das Verfahren nach der ersten Iteration, in der keine Verschiebung durchgeführt wird, da sich alle Objekte komplett in ihrer Voronoizelle befinden. Für Cluster mit mehr als 100 Objekten wird in allen Fällen die maximale Anzahl an Iterationen ausgeführt. Dies liegt daran, dass in keinem dieser Cluster eine überlappungsfreie Platzierung der Objekte gefunden wurde.

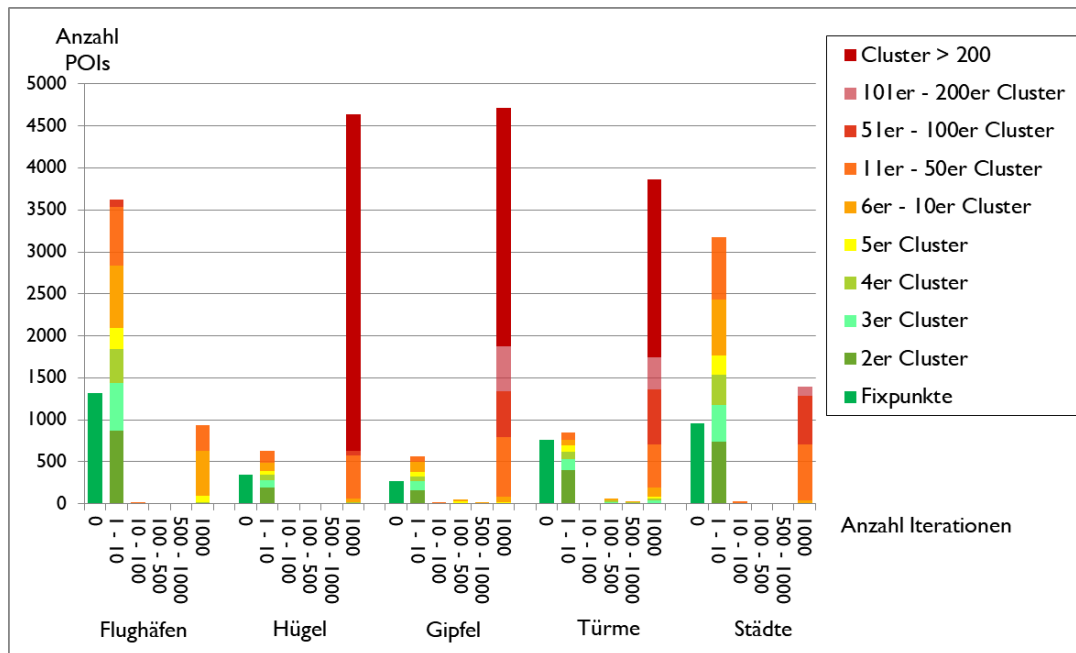


Abb. 56 Anzahl der durchgeführten Iterationen pro POI

8.2.3. Selektion zur Vermeidung unlösbarer Überlappungen

Zur Bestimmung der Anzahl der an unlösbaren Überlappungen beteiligten Signaturen werden die vier Kriterien (3er, 4er, 7er, 19er) aus Abschnitt 5.3 verwendet. Zusätzlich erfolgt noch eine Abschätzung auf Grund der Fläche die von den Signaturen überdeckt wird. Hierzu wird ein Puffer der maximalen Verschiebungsdistanz um die durch die Signaturen überdeckte Fläche konstruiert und dessen Fläche durch die Fläche einer Signatur geteilt, um zu berechnen wie viel Signaturen maximal dargestellt werden können. In Tabelle 5 ist für jede Objektart für den Datensatz 0 angegeben wie viele Signaturen an nicht überlappungsfrei darstellbaren Konstellationen beteiligt sind, wobei einige Konstellationen durch mehrere Kriterien detektiert werden. Für das Flächenkriterium ist die Anzahl der Signaturen aufgeführt, die entfernt werden müssen, um theoretisch genügend Fläche für eine überlappungsfreie Darstellung der verbliebenen Signaturen zu ermöglichen. Folglich handelt es sich hierbei um die Mindestanzahl der zu entfernenden POIs. Zusätzlich ist angegeben wie viele POIs aus den Datensätzen entfernt werden. Hierzu wird jeweils der POI, der in den meisten unlösbaren Überlappungen enthalten ist, gelöscht und anschließend überprüft, ob in der verbleibenden Punktmenge noch unlösbare Überlappungen vorhanden sind. Dieser Schritt wird solange wiederholt bis keine unlösbare Konstellation mehr gefunden wird und die Mindestanzahl erreicht ist.

Tabelle 5: Anzahl an unlösbaren Überlappungen beteiligten Signaturen

	Flughäfen		Hügel		Gipfel		Türme		Städte	
Maßstab [in Mio]	1:10	1:5	1:10	1:5	1:10	1:5	1:10	1:5	1:10	1:5
3er	0	5	3	22	22	86	14	170	0	0
4er	0	6	17	167	136	480	69	479	0	0
7er	0	7	18	187	164	594	103	509	0	0
19er	0	0	216	923	223	865	232	922	0	0
Fläche	0	0	64	495	135	586	91	619	0	0
entfernt	0	3	79	555	138	596	94	660	0	0

Bei den Stadt- und Flughafendatensätzen werden nur sehr wenige POIs entfernt, wohingegen bei den Hügel- Gipfel- und Turmdatensätze im größeren Maßstab mehr als ein Viertel der POIs entfernt werden. Bei Betrachtung der Cluster, die durch die Zerlegung der Punktmenge entstehen, ist zu sehen, dass bei den Gipfel und Turmdatensätzen keine Cluster mit mehr als 200 POIs verbleiben, bei den Hügeldatensätzen jedoch schon (siehe Abbildung 57).

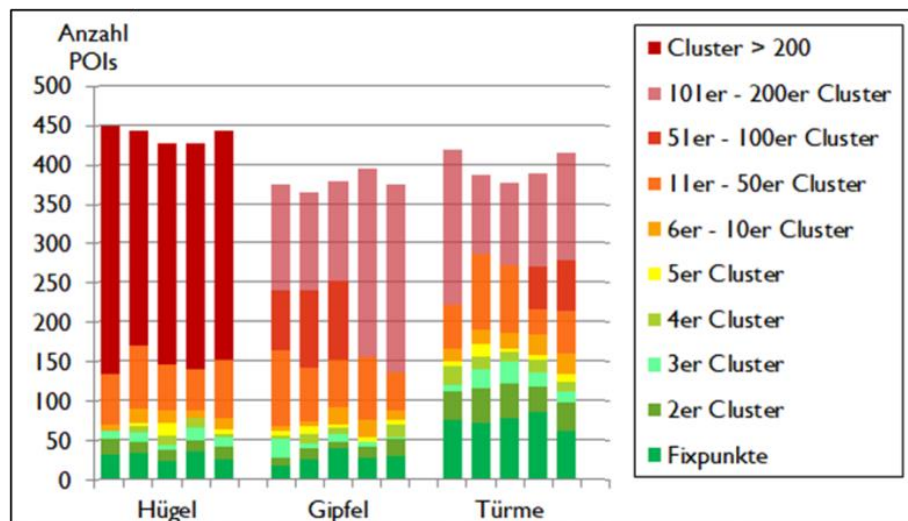


Abb. 57: Zerlegung der Datensätze in unabhängige Cluster nach der Entfernung der an unlösbaren Überlappungen beteiligten POIs

In Bezug auf die Anzahl Iterationen wird diese im gleichen Maße reduziert wie die Anzahl der POIs, da für die Cluster mit maximaler Anzahl an Iterationen auch nach der Entfernung der POIs keine überlappungsfreie Platzierung gefunden wird.

8.2.4. Lesbarkeit der Signaturen

Zur Beurteilung der Lesbarkeit der Signaturen betrachten wir jeweils die Anzahl der Signaturen, die zu weniger als der Hälfte und diejenigen, die zu weniger als drei Vierteln von anderen Signaturen überdeckt werden. Diese Daten sind in Tabelle 6 für jeweils einen der Datensätze (0) und Verwendung der Schwerpunkt-Heuristik aufgeführt. Des Weiteren ist die Anzahl der an Überlappungen beteiligten Objekte (Konflikte) angegeben.

Tabelle 6: Sichtbarkeit der Signaturen bei Anwendung der Schwerpunkt-Heuristik

		Flughäfen		Hügel		Gipfel		Türme		Städte	
Maßstab [in Mio]		1:10	1:5	1:10	1:5	1:10	1:5	1:10	1:5	1:10	1:5
sicht- bare Fläche < 1/2	vor- her	62	87	243	796	191	836	205	1357	53	249
	nach her	0	0	96	126	44	182	101	973	0	1
sicht- bare Fläche < 3/4	vor- her	110	205	303	1038	246	1025	240	1465	110	498
	nach her	0	1	171	440	95	456	140	1072	6	9
Kon- flikte	vor- her	173	438	363	1254	295	1208	269	1591	202	912
	nach her	64	132	300	996	212	984	189	1283	67	239

Bei allen Flughäfen- und Stadtdatensätzen sind nach der Verdrängung von 99% der Signaturen mehr als drei Viertel sichtbar, sowohl für die Schwerpunkt- als auch für die GIK-Heuristik. Bei den Gipfel-, Hügel- und Turmdatensätze können weniger Überlappungen aufgelöst werden und die Schwerpunkt-Heuristik führt teils zu deutlich besseren Ergebnissen, wie in Tabelle 7 zu sehen ist. Dort sind für Gipfel und Hügel im Maßstab 1:10 Mio für jeweils den Datensatz, mit den wenigsten und den meisten zu mehr als drei Vierteln überdeckten Signaturen, die Werte analog zur Tabelle 6 für beide Heuristiken aufgeführt. Die Gesamtzahl der Konflikte ist bei Verwendung der GIK-Heuristik meist geringer.

Tabelle 7 Sichtbarkeit der Signaturen vor und nach der Verdrängung bei Verwendung des Schwerpunkt-Heuristik bzw. der GIK-Heuristik

		Hügel				Gipfel			
Datensatz		2		5		0		I	
Heuristik		Schwerpunkt	GIK	Schwerpunkt	GIK	Schwerpunkt	GIK	Schwerpunkt	GIK
sichtbare Fläche < 1/2	vorher	223	223	252	252	186	186	215	215
	nachher	48	61	47	71	30	38	56	63
sichtbare Fläche < 3/4	vorher	279	279	316	316	246	246	270	270
	nachher	130	142	171	183	85	87	123	131
Konflikte	vorher	343	343	378	378	297	297	320	320
	nachher	274	276	297	281	207	198	228	218

8.2.5. Erhaltung der räumlichen Anordnung

Als quantitatives Maß für die Änderung der räumlichen Anordnung wird die Entropie der Größe der Voronoizellen (HI) verwendet. Dieses Maß ist für die gewählte Region (Europa) wenig aussagekräftig, da in den Meeren keine POIs liegen und diese somit durch deutlich größere Voronoizellen überdeckt werden, als die Landmassen. Aus diesem Grund werden die Werte für ein rechteckiges Gebiet berechnet, das in Abbildung 58 eingezeichnet ist. Im linken Bild sind die POIs analog zu Abbildung 53 dargestellt und im rechten Bild sind die POIs des Datensatzes I durch größere Punktsignaturen visualisiert.

In Abbildung 59 sind die Werte für die zehn Datensätze vor der Verdrängung pro Objekttyp und Maßstab als Boxplot-Diagramm abgebildet. Bei zehn Werten ist zwar keine statistische Auswertung möglich, dennoch bieten Boxplots eine übersichtliche Darstellungsmöglichkeit. Die HI-Werte variieren relativ stark zwischen 0,6 und 0,97, wobei eine Unterscheidung der Objekttypen an Hand des HI-Wertes nicht möglich ist. Die HI-Werte für die Flughafen- und Städtensätze sind erwartungsgemäß am größten. Bei den Gipfel Datensätzen ist der Unterschied zwischen den Datensätzen am größten. Der Grund hierfür ist, dass sich innerhalb des gewählten Gebietes ein großer Cluster (Alpen) befindet und nur sehr wenige vereinzelte Gipfel außerhalb des Clusters, so dass der HI-Wert sich relativ stark ändert je nachdem wie viele der einzelnen Gipfel enthalten sind.

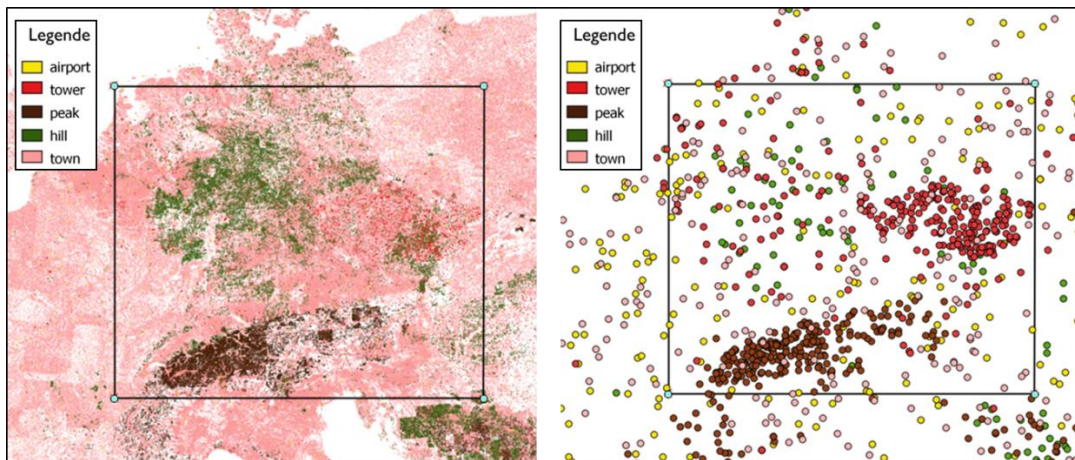


Abb. 58: Wahl des Gebietes zur Berechnung der HI-Werte

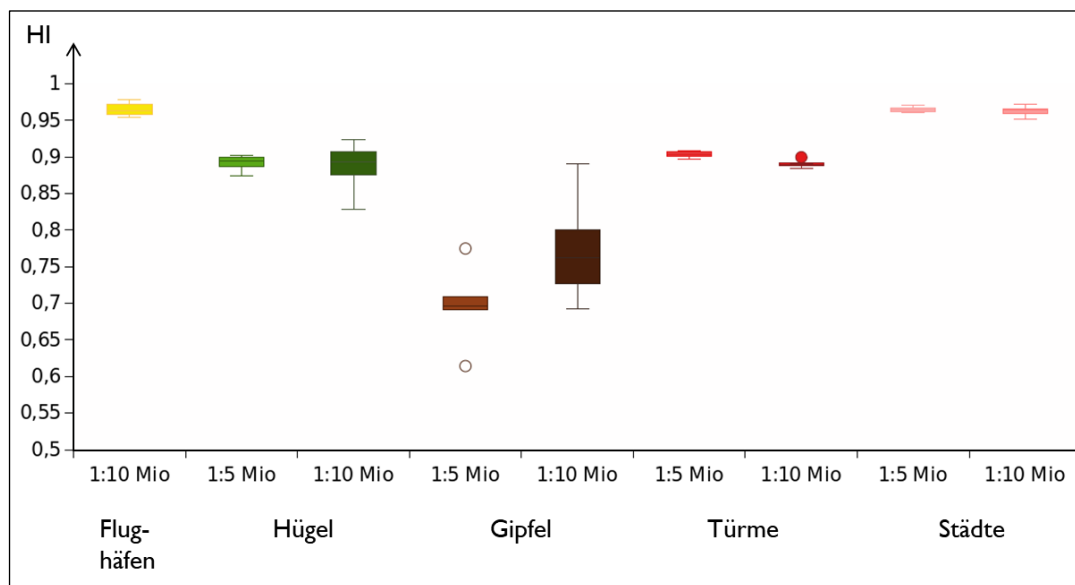


Abb. 59 HI-Werte der Datensätze vor der Verdrängung

Die Differenz der HI-Werte der Datensätze vor und nach der Verdrängung liegt in allen Fällen unter 0,05, wie in Abbildung 60 zu sehen ist. Insbesondere für die Flughäfen und Hügel Datensätze sind die Änderungen mit weniger als 0,02 sehr gering. Die Änderung des HI-Wertes durch die Verdrängung mit der Schwerpunkt-Heuristik (Centroid) ist etwas größer als bei Anwendung der GIK-Heuristik und bei den Datensätzen des kleineren Maßstabs ist die Änderung größer als bei den Datensätzen des größeren Maßstabs. Zur Veranschaulichung der Bedeutung der Werteänderung sind in Abbildung 61 für die Hügel Daten, der Datensatz mit dem kleinsten HI-Wert (linke Spalte), dem größten HI-Wert (mittlere Spalte) und der größten HI-Wertänderung (rechte) Spalte, jeweils vor der Verdrängung (obere

Reihe), nach Verdrängung mit der GIK-Heuristik (mittlere Reihe) und nach der Verdrängung mit der Schwerpunkt-Heuristik (untere Reihe) dargestellt.

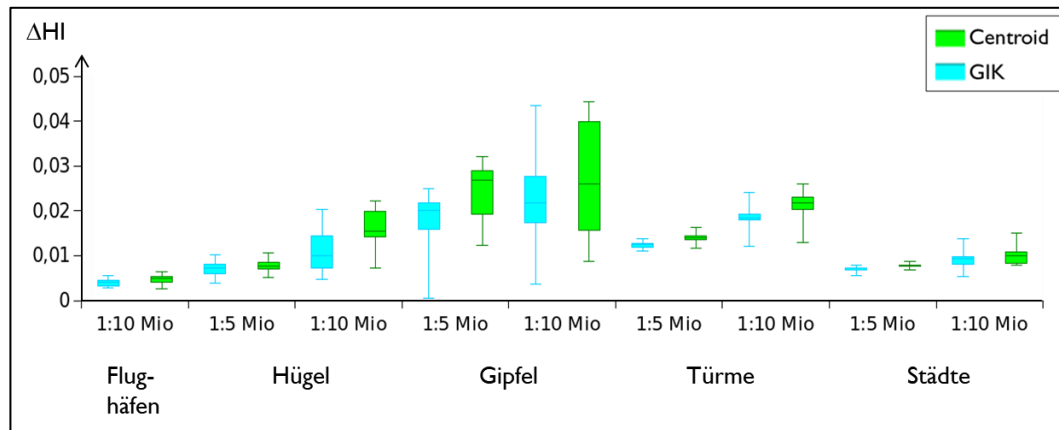


Abb. 60: Vergleich der Änderung der HI-Werte bei Verwendung der GIK- und Schwerpunkt-Heuristik

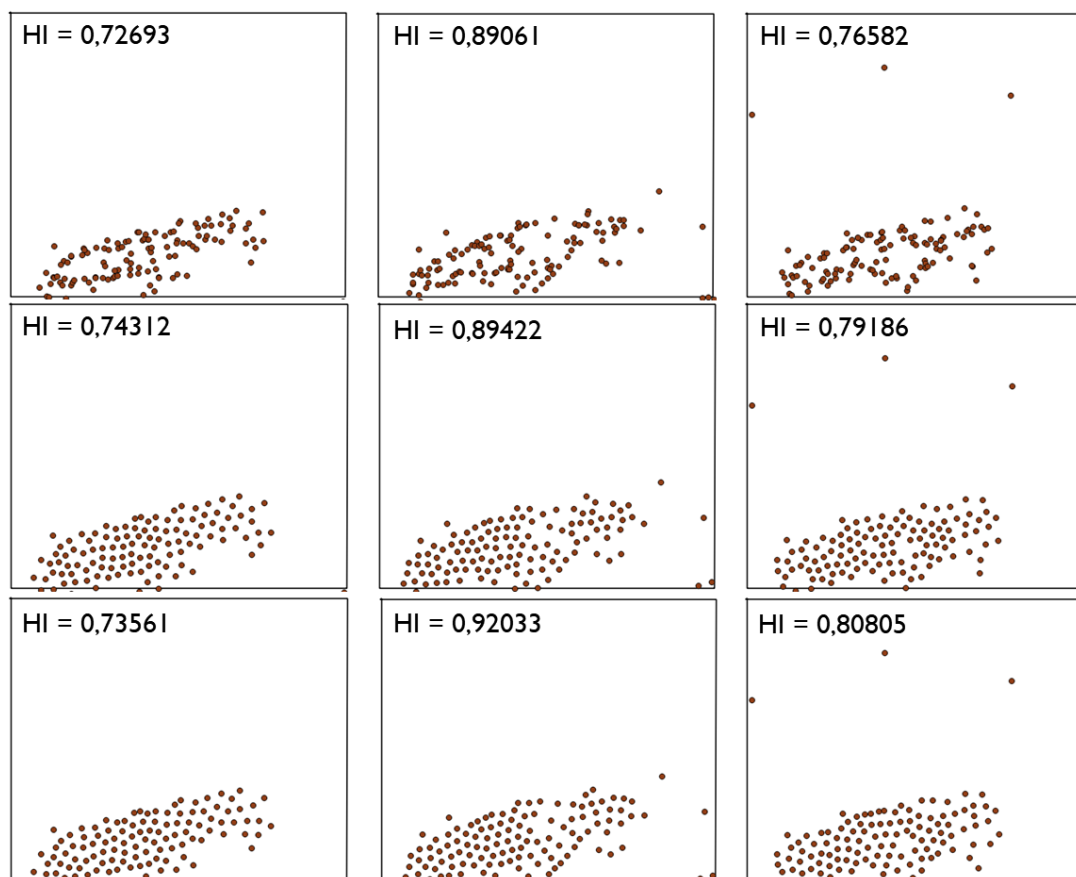


Abb. 61: Verteilung der Gipfelsignaturen vor und nach der Verdrängung

Zusätzlich wurde KDE angewendet, um die Beibehaltung der räumlichen Anordnung der Punkte qualitativ zu vergleichen. In Abbildung 62 sind jeweils Ausschnitte der Kernel-Dichte-Karten für den Flughafendatensatz I vor (a) und nach (b) der Verdrängung mit der Schwerpunkt-Heuristik, wie auch ein Teil des Gipfel-datensatzes I in (c) und (d), jeweils im kleineren Maßstab, abgebildet.

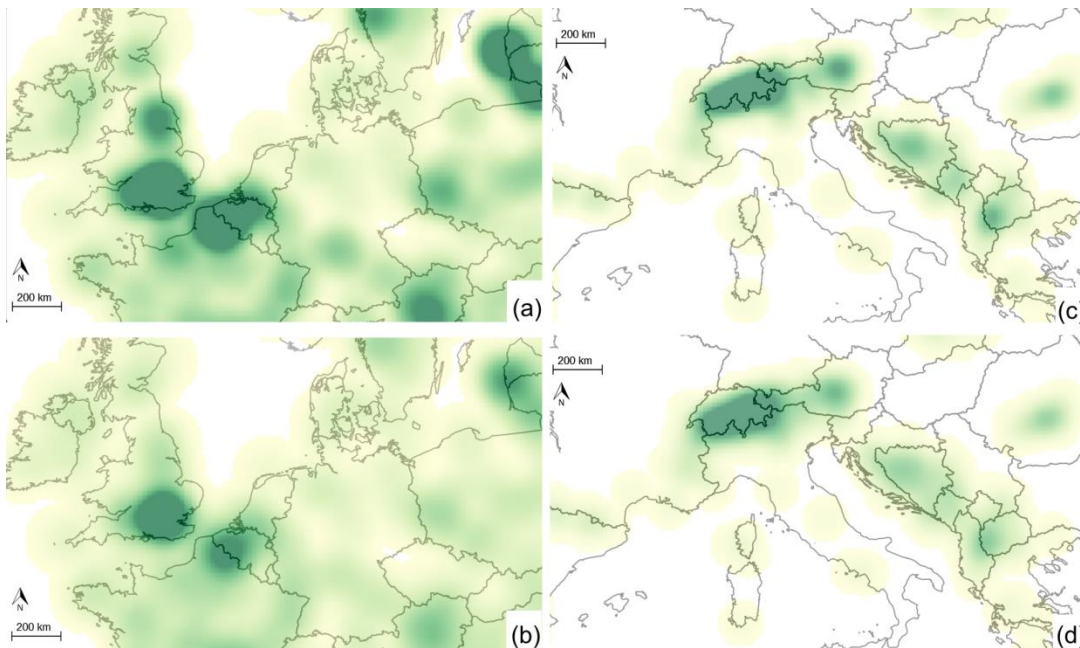


Abb. 62: Kernel density maps: Flughäfen/Gipfel vor (a/c) und nach (b/d) der Verdrängung

Nach der Verdrängung wirken die Gebiete, in denen sich die lokalen Dichte-Maxima befinden, verwischt, da ihre Ausdehnung vergrößert und die Maximalwerte verringert wurden. Dennoch bleibt das zugrundeliegende räumliche Muster erhalten. Ein qualitativer Unterschied bei den Ergebnissen für die beiden Heuristiken ist nicht feststellbar.

8.2.6. Fazit

In Bezug auf die Erhaltung der räumlichen Anordnung erfüllen die Ergebnisse aller Datensätze die in dieser Arbeit gestellten Anforderungen, da die globale Verteilung erhalten bleibt, auch wenn in Gebieten mit vielen Überlappungen die Signaturen auf eine größere Fläche verteilt werden und sich so lokal eine Änderung der Anordnung ergibt. Für die Flughäfen- und Städtedatensätze wird eine deutlich bessere Lesbarkeit der einzelnen Signaturen erreicht, da nach dem Verdrängungsverfahren mindestens 99% der Signaturen, die zuvor stärker verdeckt waren, zu mehr als drei Viertel sichtbar sind. Bei den Hügel-, Gipfel- und Turm-datensätzen wird nur

für 25% bis zu 40% der Signaturen eine signifikante Verbesserung der Lesbarkeit erreicht. Der Grund hierfür ist der hohe Anteil an unlösbaren Überlappungen, der durch die untere Schranke (siehe Abschnitt 5.3) auf über 50% berechnet wird. Die Anzahl der notwendigen Iterationen ist ebenfalls abhängig von den unlösbaren Überlappungen. Für Teilmengen, für die überlappungsfreie Anordnungen gefunden werden, werden in über 99% weniger als 10 Iterationen und maximal 74 Iterationen durchgeführt, für Teilmengen mit unlösbaren Überlappungen die maximale Anzahl an Iterationen. Beim Vergleich der beiden Heuristiken schneidet die Schwerpunkt-Heuristik in den meisten Fällen, sowohl in Bezug auf die Erhaltung der räumlichen Verteilung als auch die Sichtbarkeit der Signaturen, etwas besser ab.

In Abbildung 63 sind die Flughäfen des dritten Datensatzes (1:10 Mio) links vor der Verdrängung und rechts nach der Verdrängung in einer kartenähnlichen Grafik nicht maßstabstreu dargestellt. Im linken Bild sind disjunkte und überlappende POIs durch Signaturen mit verschiedenen Farben symbolisiert und im rechten wird farblich zwischen POIs, die während der Verdrängung nicht berücksichtigt werden (Fixpunkte), POIs die nicht verdrängt wurden, POIs die verdrängt wurden und sich nicht mehr überlappen sowie Überlappungen die bestehen bleiben, unterschieden. Analog sind in Abbildung 64 die Städte des sechsten Datensatzes (1:5 Mio) vor der Verdrängung und in Abbildung 65 nach der Verdrängung dargestellt.

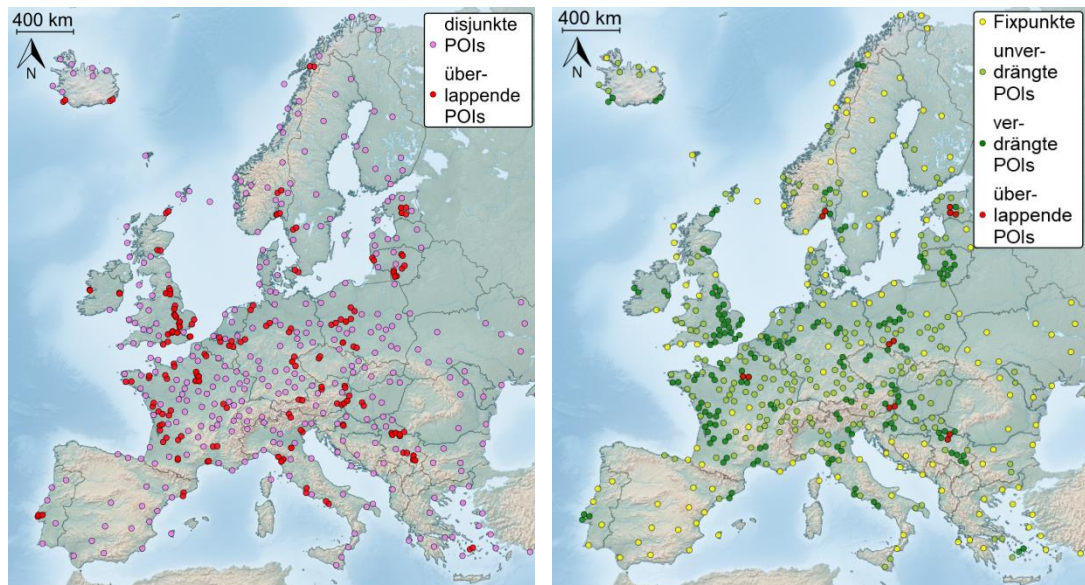


Abb. 63: Flughäfen vor (links) und nach der Verdrängung (rechts)



Abb. 64: Städte vor der Verdrängung



Abb. 65: Städte nach der Verdrängung

8.3. Vergleich zwischen Verwendung des Power- und des Voronoi-Diagramms

Als Testdatensätze werden die Stadt und die Gipfel datensätze aus Abschnitt 8.2 verwendet, wobei zusätzlich zur Berechnung der Größe der Signatur die Höhe der Gipfel und die Einwohnerzahl der Städte berücksichtigt werden. Bei den Städten werden nur Städte mit mehr als 1000 und weniger als 10000 Einwohner ausgewählt. Da andernfalls bei der Normierung der Werte auf das Intervall $[0;1]$ 99% der Daten einen Wert erhalten der kleiner als 0,1 ist. In Abbildung 66 ist die Werteverteilung dargestellt.

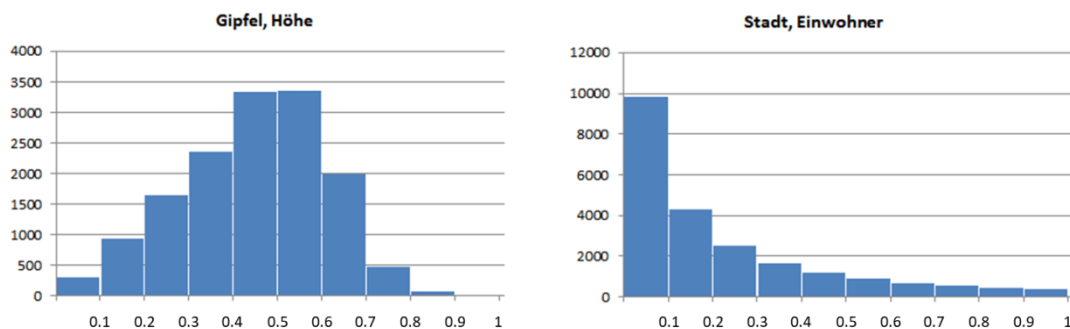


Abb. 66: Normierte Werteverteilung der Gipfelhöhen (links) und Einwohnerzahlen (rechts)

Zur Berechnung der Signaturgröße wird für jeden POI der Radius r , der in Abschnitt 8.2 verwendet wird, mit dem normierten Wert multipliziert und anschließend $r/2$ addiert. Dies bedeutet, dass die Werte der Radien im Intervall $[0,5r; 1,5r]$ liegen. In Abbildung 67 ist ein Ausschnitt des Stadt datensatzes 0 vor der Verdrängung dargestellt.

Für die Datensätze 0-4 wird das Verdrängungsverfahren jeweils für zwei Werte (r , $1,5r$) für die maximale Verschiebungsdistanz durchgeführt, wobei einmal das Voronoi-Diagramm (V) und einmal das Power-Diagramm (P) als Hilfsstruktur verwendet wird. In Tabelle 8 sind die Ergebnisse für den Datensatz 0 für die Städte und Gipfel analog zur Tabelle 6 für die Schwerpunkt-Heuristik und dem Maßstab 1:10 Mio aufgelistet.

Bei den Stadt datensätzen reduziert sich die Anzahl der Konflikte bei Verwendung des Power-Diagramms stark, so wird sie bei der Variante mit dem Voronoi-Diagramm von 224 auf 169 reduziert und durch die Power-Diagramm-Variante von 224 auf 2. Für die Gipfel datensätze wird die Anzahl der Konflikte insgesamt verringert, die Anzahl der Signaturen, die nach der Verdrängung zur Hälfte bzw. zu drei Vierteln sichtbar sind, wird ebenfalls verringert. Dies könnte darin begründet sein, dass bei der Verwendung des Powerdiagramms als Hilfsstruktur, tendenziell die größeren Signaturen verschoben werden, wie z.B. beim Vergleich der Abbildungen 68 und 69 in Luxemburg zu sehen ist. In Abbildung 68 ist das Ergebnis der Verdrängung der in Abbildung 67 dargestellten Signaturen mittels Voronoi-

Diagramm dargestellt und in Abbildung 69 die mit Hilfe des Power-Diagramms berechnete Platzierung.

Tabelle 8: Sichtbarkeit unterschiedlich großer Signaturen vor und nach der Verdrängung bei Verwendung des Voronoi-Diagramms (V) bzw. des Power-Diagramms (P)

		Städte				Gipfel			
Verschiebungsdistanz		r		1,5r		r		1,5r	
Verfahren		V	P	V	P	V	P	V	P
sichtbare Fläche < 1/2	vorher	126	126	126	126	294	294	294	294
	nachher	4	0	4	0	207	213	207	215
sichtbare Fläche < 3/4	vorher	44	44	44	44	264	264	264	264
	nachher	0	0	1	0	163	178	164	178
Konflikte	vorher	224	224	224	224	356	356	356	356
	nachher	169	2	164	2	344	260	338	260

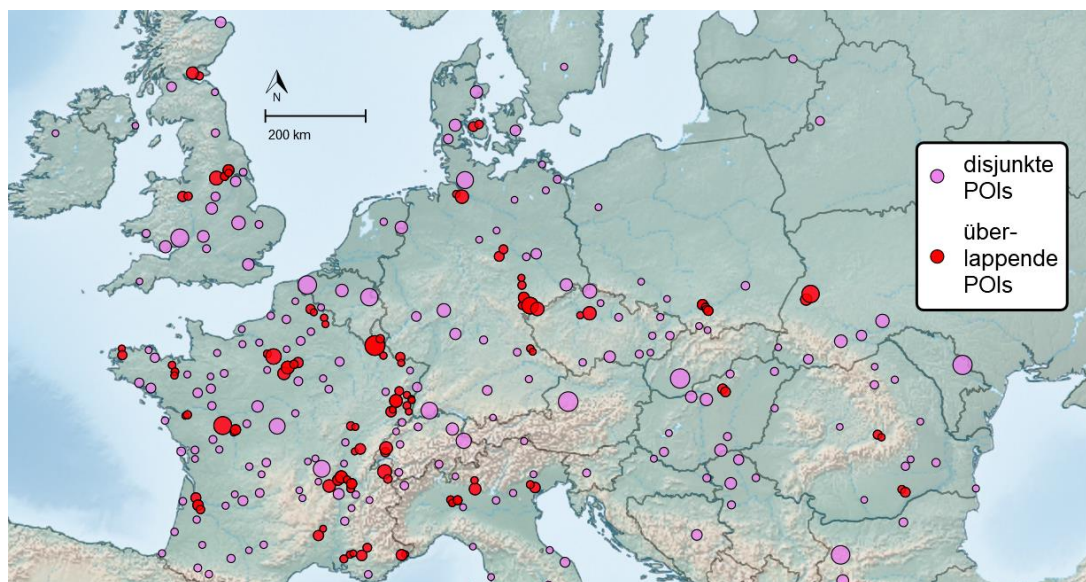


Abb. 67: Unterschiedliche große Stadtsignaturen vor der Verdrängung

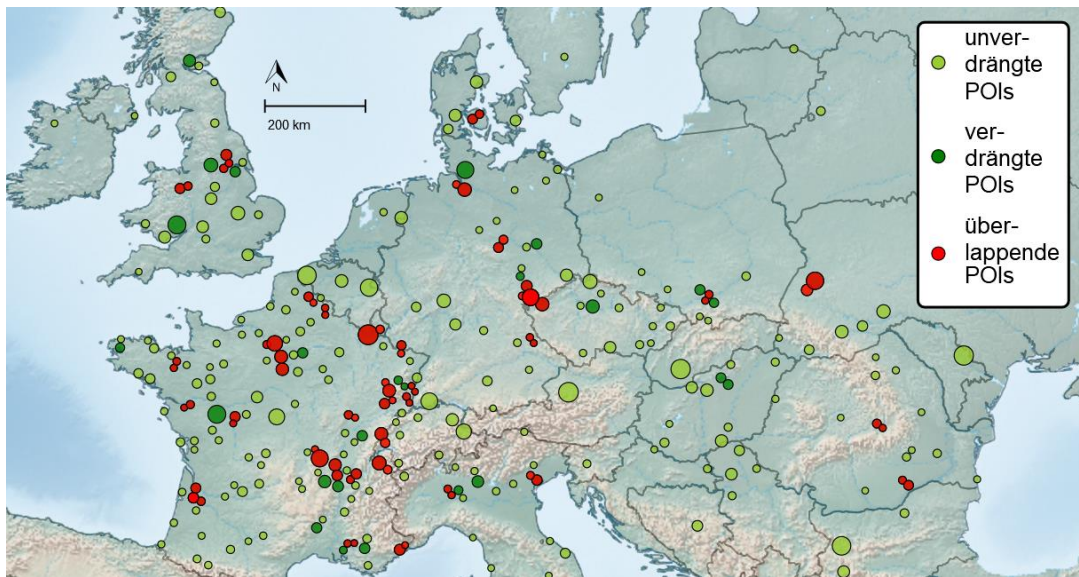


Abb. 68: Unterschiedlich große Stadtsignaturen nach der Verdrängung mittels Voronoi-Diagramm

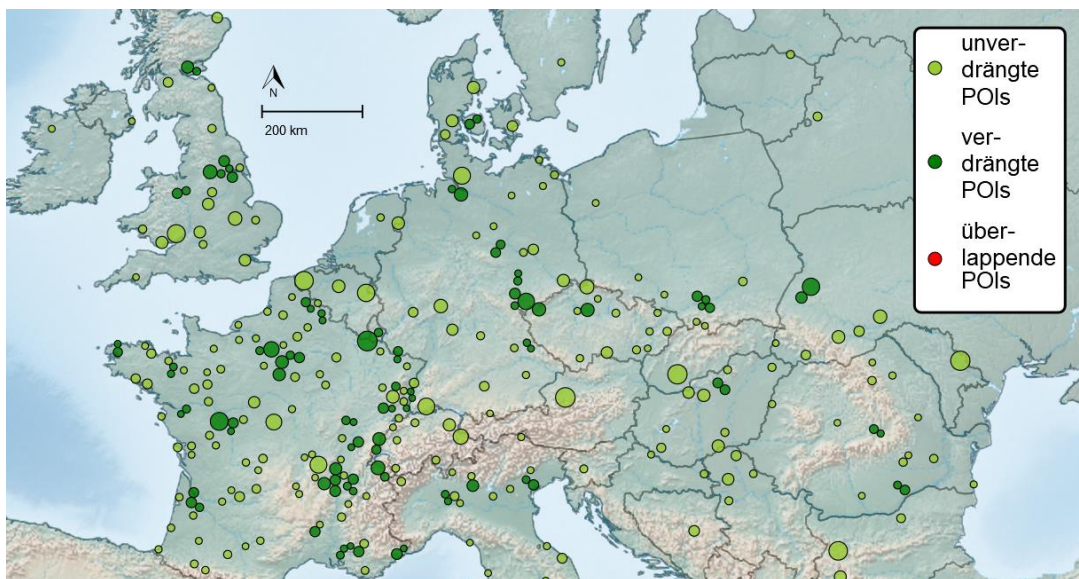


Abb. 69: Unterschiedlich große Stadtsignaturen nach der Verdrängung mittels Power-Diagramm

Eine Verbesserung der Sichtbarkeit wird durch die Erhöhung der maximalen Verschiebungsdistanz nicht erreicht. In einigen Fällen fallen die Ergebnisse sogar geringfügig schlechter aus.

8.4. Anpassen der Zellgeometrie

Als Testdatensätze werden erneut die Datensätze aus Abschnitt 8.3 verwendet, wobei die normierten Einwohnerzahlen und Gipfelhöhen für positive Werte verwendet werden und für die negativen Werte eine Pseudo-Zufallszahl im Intervall $[0;1]$ generiert wird. Die Berechnung des Radius erfolgt dann wie beim positiven Wert. In Abbildung 70 (mitte) ist ein Ausschnitt des Gipfel Datensatzes vor der Verdrängung visualisiert, bei der die Halbkreisignatur aus Abschnitt 6.4 verwendet wird. Wie in Abschnitt 8.3 wird für die Datensätze 0 bis 4 das Verdrängungsverfahren jeweils für zwei Werte (r , $1,5r$) für die maximale Verschiebungsdistanz durchgeführt. Als Hilfsstruktur wird jeweils das Power-Diagramm verwendet; einmal mit und einmal ohne Anpassung der Zellgeometrie, wie in Abschnitt 7.2.2 beschrieben. In Tabelle 9 sind die Ergebnisse für den Datensatz 0 für die Städte und Gipfel analog zur Tabelle 6 für die Schwerpunkt-Heuristik und dem Maßstab 1:10 Mio aufgelistet. Zur Bestimmung der Sichtbarkeit wird für jeden Halbkreis separat berechnet, wie groß der sichtbare Anteil ist, und das Minimum wird für die Berechnung der Angaben in der Tabelle verwendet.

Die Ergebnisse ähneln denen aus Abschnitt 8.3: Eine Vergrößerung der Verschiebungsdistanz führt nicht zu einer Verbesserung der Ergebnisse und durch die Variante, bei der die Signaturform stärker berücksichtigt wird, werden mehr Konflikte aufgelöst, die Sichtbarkeit ist jedoch schlechter. Anders als bei der Variante aus Abschnitt 8.3 ist jedoch kein qualitativer Unterschied zwischen den Ergebnissen zu sehen, auch wenn sich die Platzierung der Punkte teils deutlich unterscheidet, wie z.B. in Abbildung 71 im pink umrandeten Gebiet.

Tabelle 9: Sichtbarkeit unterschiedlich großer Signaturen vor und nach der Verdrängung bei Verwendung des Power-Diagramms ohne (P) und mit der Anpassung der Zellgeometrie (AZ)

		Städte				Gipfel			
Verschiebungsdistanz		r		1,5r		r		1,5r	
Verfahren		P	AZ	P	AZ	P	AZ	P	AZ
sichtbare Fläche < 1/2	vorher	94	94	94	94	292	292	292	292
	nachher	2	0	2	0	159	186	161	186
sichtbare Fläche < 3/4	vorher	72	72	72	72	270	270	270	270
	nachher	1	0	1	0	138	161	137	162
Konflikte	vorher	308	308	308	308	394	394	394	394
	nachher	156	138	153	142	344	335	343	340

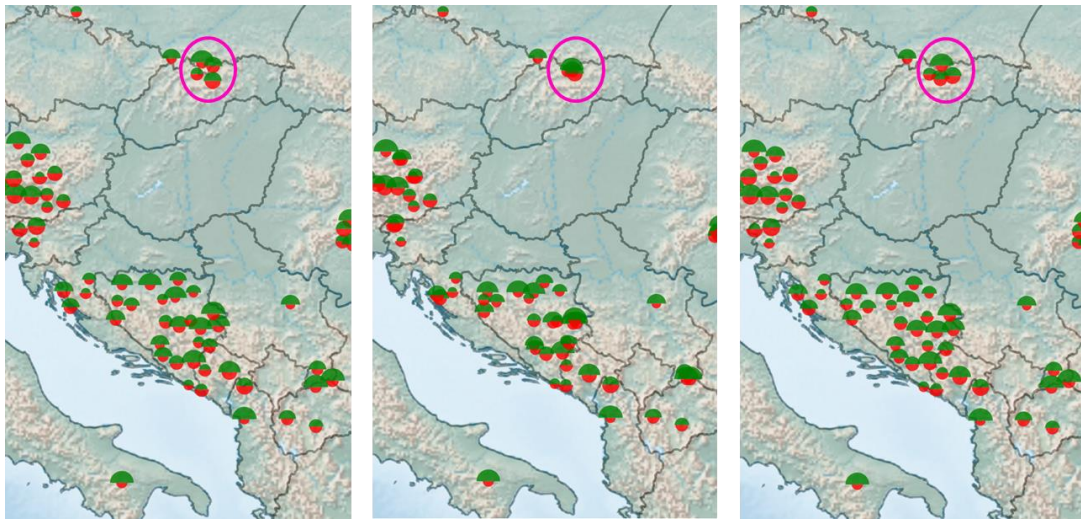


Abb. 70: Unterschiedlich große Gipfelsignaturen vor (mitte), nach der Verdrängung mittels Power-Diagramm ohne (links) und mit Anpassung der Zellgeometrie (rechts)

8.5. Abschätzung des Signaturradius

Wie in Abschnitt 7.1.4 erläutert ist es möglich nach der Durchführung des Verdrängungsverfahrens für die Punktmenge den Radius r_v zu schätzen, für den eine überlappungsfreie Darstellung theoretisch möglich ist, wenn die Punktsignaturen gleich groß sind. Basierend auf den Ergebnissen aus Abschnitt 8.2.4 wird für jeden der Datensätze der Radius r_v berechnet und anschließend erneut das Verdrängungsverfahren durchgeführt, wobei die maximale Verschiebungsdistanz wie in Abschnitt 8.2.4 2mm ist. In Tabelle 10, sind für den Datensatz 0 die Anzahl der Signaturen, die zu mehr als drei Viertel sichtbar sind und die Anzahl der Konflikte jeweils vor und nach der Verdrängung angegeben. Die Anzahl der zur Hälfte sichtbaren Signaturen ist nach der Verdrängung für alle Datensätze 0 oder 1 und nicht in der Tabelle aufgeführt. Zusätzlich ist noch der Radius r_v angegeben und die durchschnittliche Distanz um die eine überlappte Signatur verschoben wird. Hierzu wird die Summe der Länge aller Verschiebungsvektoren durch die Anzahl der Konflikte geteilt und in Relation zu r_v gesetzt. So werden Signaturen der Flughafen- und Stadtdatensätze im Schnitt ca. um die Hälfte des Radius r_v verschoben. Die Verschiebungsdistanz ist bei den Hügel-, Gipfel- und Turmdatensätze mit 83% - 94% deutlich höher, jedoch immer geringer als der Radius r_v , obwohl dieser für die entsprechenden Datensätze deutlich geringer (25% bis 35%) als die maximale Verschiebungsdistanz ist. In Bezug auf die Sichtbarkeit sind bei allen Stadt- und Flughafendatensätzen alle Signaturen zu mehr als drei Viertel sichtbar und für die anderen Datensätze werden nach der Verdrängung maximal 1,6% der Signaturen zu mehr als drei Viertel überdeckt.

Tabelle 10: Sichtbarkeit der Signaturen nach Änderung der Signaturreadien

		Flughäfen		Hügel		Gipfel		Türme		Städte	
Maßstab		1:10	1:5	1:10	1:5	1:10	1:5	1:10	1:5	1:10	1:5
sichtbare Fläche < 3/4	vorher	102	168	226	733	177	743	181	593	91	370
	nachher	0	0	0	0	1	10	0	8	2	2
Konflikte	vorher	160	356	309	1039	267	1014	237	777	172	714
	nachher	40	45	74	159	74	157	65	108	25	83
Radius r_v (mm)		1,9	1,8	1,4	1,4	1,5	1,3	1,3	1,3	1,8	1,7
Rel. Verschiebungsdistanz		0,54	0,44	0,92	0,87	0,83	0,91	0,94	0,88	0,47	0,50

In Abbildung 72 ist der Hügel datensatz 5 vor der Verdrängung und nach der in Abschnitt 8.4 durchgeführten Verdrängung abgebildet. Basierend auf dem Ergebnis wurde neben den Radius r_v analog der Radius r_h berechnet, für den theoretisch eine Anordnung möglich ist, bei der jede Signatur mindestens zur Hälfte sichtbar ist. In Abbildung 73 ist der Hügel datensatz vor und nach der Verdrängung mit Signaturen mit Radius r_h und in Abbildung 73 für Signaturen mit dem Radius r_v dargestellt.

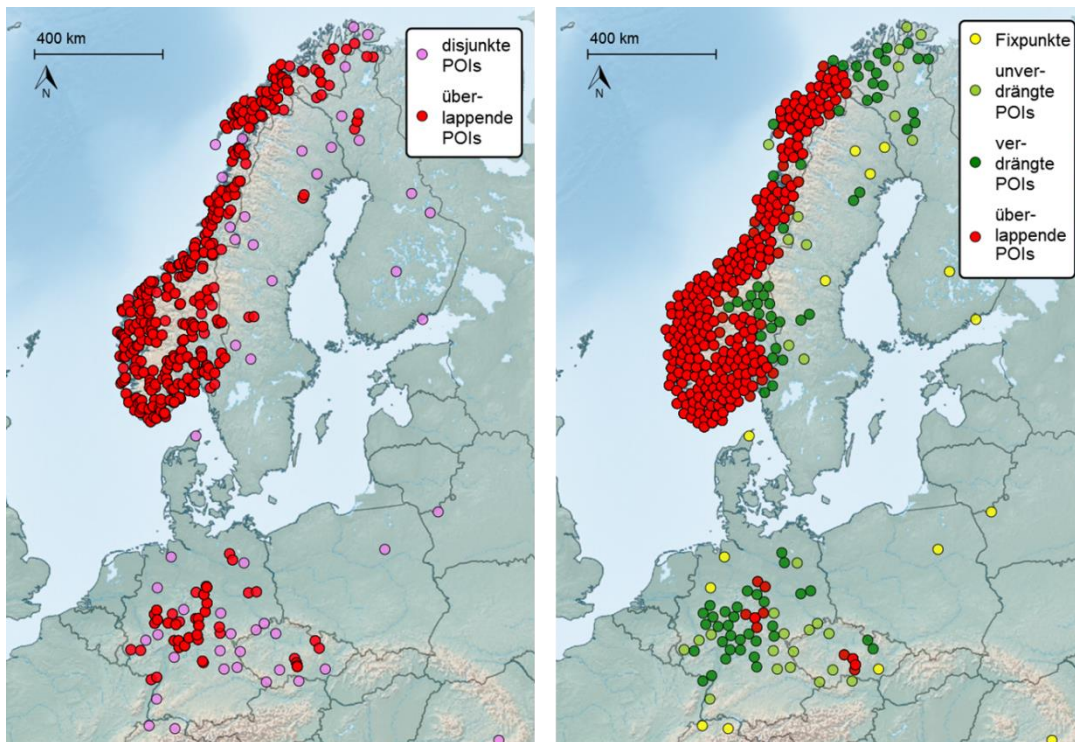


Abb. 71: Hügel signatures mit Radius r vor (links) und nach der Verdrängung (rechts)

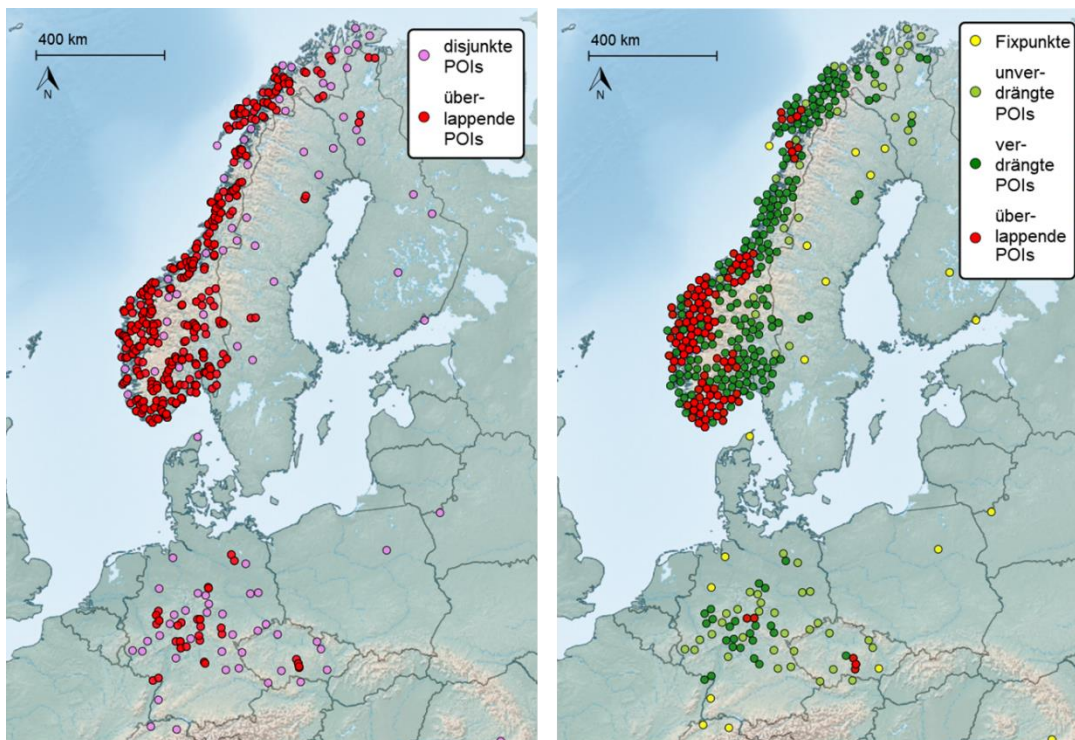


Abb. 72: Hügel mit Radius r_h vor (links) und nach der Verdrängung (rechts)

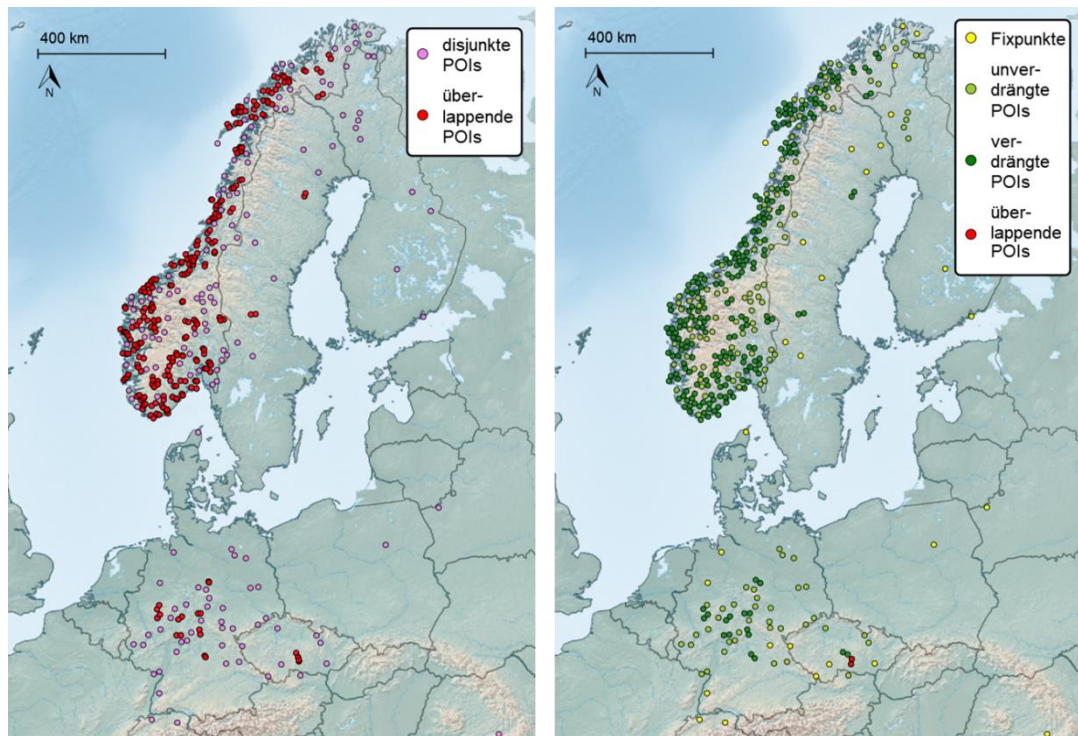


Abb. 73: Hügel mit Radius r_v vor (links) und nach der Verdrängung (rechts)

Zusammenfassung und Ausblick

Im Rahmen dieser Arbeit wird die Visualisierung von impliziten geographischen Informationen an Hand eines konkreten Anwendungsfalls, nämlich der Visualisierung von Meinungen und Gefühlen (Sentiments) mit Ortsbezug betrachtet. Hierzu können drei verschiedene Ansätze verfolgt werden:

- 1) Darstellung durch Anpassung vorhandener Signaturen
- 2) Darstellung als Kontinua
- 3) Darstellung durch Punktsignaturen

Der dritte Ansatz bietet die größte Flexibilität und kann auch auf komplexere Sentimentmodelle, als das in dieser Arbeit betrachtete eindimensionale Modell (Polarität), erweitert werden. Allerdings ist es für eine lesbare Darstellung wichtig, dass sich die Symbole nicht überlappen, insbesondere wenn die verwendeten Signaturen nicht identisch sind, sondern über ihre graphischen Variablen weitere Informationen visualisiert werden sollen. Aus diesem Grund sind, für die Realisierung eines Systems, das den Benutzer bei der Generierung lesbarer Karten unterstützt, Generalisierungsverfahren zur Auflösung von Überlappungen von Punktsignaturen notwendig.

Es wird ein iteratives Verfahren zur Verdrängung von Punktsignaturen vorgestellt und evaluiert. In Bezug auf die Erhaltung der räumlichen Anordnung bleibt die globale Verteilung erhalten, auch wenn in Gebieten mit vielen Überlappungen die Signaturen auf eine größere Fläche verteilt werden und sich so lokal eine Änderung der Anordnung ergibt. Die Verbesserung der Lesbarkeit ist abhängig von der Anzahl an nicht auflösbaren Überlappungen. In Fällen in denen keine überlappungsfreie Darstellung möglich ist, werden durch das Verfahren die Signaturen so angeordnet, dass sich mehr Signaturen überlappen, aber die Überlappungsfläche für jede Signatur möglichst gering ist. Aus diesem Grund ist ein Kriterium wichtig, um abschätzen zu können, ob eine lesbare Darstellung einer Punktmenge überhaupt erreicht werden kann. Hierfür wurden Konstellationen von POIs identifiziert, die nicht überlappungsfrei darstellbar sind, wenn der Signaturreadius fix und gleich der maximalen Verschiebungsdistanz ist. Auf diese Weise lässt sich für eine gegebene POI-Menge eine untere Schranke für die Anzahl der Konflikte angeben, die nicht aufgelöst werden können. Wenn die Signaturgröße variabel ist, wurde eine Möglichkeit aufgezeigt einen Signaturreadius zu berechnen, der eine weitgehend überlappungsfreie Darstellung ermöglicht. Hierzu wird zunächst, das Verdrängungsverfahren mit einem vorgegebenen Radius durchgeführt und auf Grundlage des Ergebnisses ein geeigneter Signaturreadius berechnet. Anschließend wird das Verfahren

erneut mit diesem Radius und gleicher maximaler Verschiebungsdistanz durchgeführt.

In Bezug auf die Parameter des Verdrängungsverfahrens, ist die Wahl des Signaturradius relevanter, als die der Verschiebungsdistanz, um eine gute Visualisierung zu erreichen. In den in dieser Arbeit durchgeführten Experimenten, hat eine realistische Vergrößerung der maximalen Verschiebungsdistanz keine nennenswerte Verbesserung der Ergebnisse zur Folge. Dagegen führt eine Verkleinerung des Radius neben der Reduzierung der Überlappungen auch zu einer Verringerung der Distanz, um die die Signaturen verdrängt werden, ohne dass die maximale Verschiebungsdistanz verkleinert wird.

Die Anzahl der Iterationen ist ebenfalls davon abhängig, ob eine überlappungsfreie Darstellung gefunden werden kann. In Fällen in denen dies nicht der Fall ist, wird die maximale Anzahl an Iterationsschritten ausgeführt. Um dieses zu vermeiden, wäre es möglich ein Abbruchkriterium, wie z.B. die Gesamtüberlappungsfläche oder Anzahl der Konflikte zu definieren und das Verfahren zu beenden, wenn keine Verbesserung mehr eintritt, oder ein hinreichend gutes Ergebnis erreicht wurde.

Bezüglich der Heuristiken konnten im Rahmen der Experimente nur geringe Unterschiede festgestellt werden, jedoch erscheinen weitere Untersuchungen im Hinblick auf die Eignung der Heuristiken für verschiedene Punktmengen und auch in Bezug auf die Varianten des Verdrängungsverfahrens (Power-Diagramm, Anpassung der Zellgeometrie) lohnenswert. Hierzu wären zunächst Kriterien zu definieren um die Qualität verschiedener überlappungsfreier Anordnungen zu bewerten.

Abbildungsverzeichnis

ABB. 1: DARSTELLUNG EINER ROUTE IN GOOGLE MAPS.....	4
ABB. 2: BEISPIEL EINER NUTZERGENERIERTEN KARTE MITTELS GOOGLE MAPS	5
ABB. 3: VERSCHIEDENE VARIANTEN VON VORONOI-DIAGRAMMEN	9
ABB. 4: BEISPIELE FÜR SCHWERPUNKT-VORONOI-DIAGRAMME	11
ABB. 5: KONSTRUKTION VON KANTEN DES POWER-DIAGRAMMS.....	12
ABB. 6: BEISPIEL EINES POWER-DIAGRAMMS	13
ABB. 7: BEISPIELE FÜR FORMEN UND ANORDNUNGEN VON SIGNATUREN[HGM02].....	16
ABB. 8: GRAPHISCHE VARIABLEN	17
ABB. 9: EINFLUSS DES HINTERGRUNDES AUF WAHRNEHMUNG DES HELLWERTS	19
ABB. 10 GESICHTSAUSDRÜCKE ZUR VISUALISIERUNG VON BASISEMOTIONEN UND GEMISCHTEN EMOTIONEN [MC06].....	28
ABB. 11 BODILY MAPS VON EMOTIONEN [NGH+13].....	28
ABB. 12 COMFORT MAPS: EINZELDARSTELLUNG UND AGGREGIERTE DARSTELLUNG [MBQ10]	29
ABB. 13 AUSSCHNITTE AUS KARTEN DES BIOMAPPING-PROJEKTS [NO09]	30
ABB. 14 VISUALISIERUNG VON SENTIMENTS DURCH SMILIES IN VERSCHIEDENEN ZOOMSTUFEN [CDG+14].....	31
ABB. 15: TOPOLOGISCH DIFFERENZIERBARE ÜBERLAPPENDE ANORDNUNGEN VON 4 TRANSPARENTEN KREISEN	39
ABB. 16: TOPOLOGISCH DIFFERENZIERBARE ÜBERLAPPENDE ANORDNUNGEN VON 4 SIGNATUREN	39
ABB. 17: DETEKTION VON ÜBERLAPPUNGEN MIT BELIEBIGER UND DELAUNAY- TRIANGULATION	40
ABB. 18: INNERER UND ÄUßERER SODDY-KREIS.....	41
ABB. 19: DICHTESTE PACKUNGEN VON 4, 7 UND 19 KREISEN IN EINEM KREIS	42
ABB. 20: EINFLUSS DER HINTERGRUNDKARTE AUF DEN SIGNATURENTWURF	43
ABB. 21 AUSCHNITT DER LEGENDE DER TK25	46
ABB. 22 VERSCHIEDENE WERTKOMBINATIONEN BEI KODIERUNG DER SENTIMENTSTÄRKE DURCH GRÖÖE UND VON “POSITIV/NEGATIV” DURCH RICHTUNG (LINKS) UND FORM (RECHTS)	47
ABB. 23: VERSCHIEDENE WERTKOMBINATIONEN BEI KODIERUNG DER SENTIMENTSTÄRKE DURCH IHREN HELLWERT UND VON “POSITIV/NEGATIV” DURCH RICHTUNG (LINKS) UND FORM (RECHTS)	48
ABB. 24: VARIANTEN ZUR ANPASSUNG VON FLÄCHENSIGNATUREN MIT UNTERSCHIEDLICHEM HELLWERT	49
ABB. 25 VARIANTEN ZUR ANPASSUNG VON FLÄCHENSIGNATUREN MIT GLEICHEM HELLWERT	49
ABB. 26: POSITIVE UND NEGATIVE SENTIMENTWERTE	50
ABB. 27: KOMBINATION POSITIVER UND NEGATIVER SENTIMENTWERTE.....	51
ABB. 28: VERSCHIEDENE FARBSCHEMATA FÜR SENTIMENTS.....	52
ABB. 29: SENTIMENTKARTE FÜR REYKJAVIK	52
ABB. 30: KOMBINIERT PUNKTSIGNATUREN FÜR UNTERSCHIEDLICHE SENTIMENTWERTE	54

ABB. 31: KONSTRUKTION DER STERNSIGNATUR	55
ABB. 32: ÜBERDECKUNG RELEVANTER HINTERGRUNDOBJEKTE DURCH SIGNATUREN	57
ABB. 33: ANFANGSPLATZIERUNGEN VON POIs IN VORONOIZELLEN: MÖGLICHE KONSTELLATIONEN.....	58
ABB. 34: ITERATION 1 BIS 3 UND ERGEBNIS DURCH HEURISTIK A).....	59
ABB. 35: ITERATION 1 BIS 3 UND ERGEBNIS DURCH HEURISTIK B)	59
ABB. 36: VERGRÖßERUNG DES ABSTANDS ZUM NÄCHSTEN NACHBARN DURCH VERSCHIEBUNG IN RICHTUNG DES MITTELPUNKTS DES GIK.....	61
ABB. 37: ÄNDERUNG DER GESAMTÜBERLAPPUNGSFLÄCHE DER SIGNATUREN.....	62
ABB. 38: GRÖßE DER ÜBERLAPPUNGSFLÄCHE IN ABHÄNGIGKEIT DES ABSTANDS.....	63
ABB. 39: VERGRÖßERUNG DER ÜBERLAPPUNGSFLÄCHE DURCH VERSCHIEBUNG AUF DEN SCHWERPUNKT.....	64
ABB. 40: ÜBERWINDUNG EINES LOKALEN MINIMUMS DURCH DIE SCHWERPUNKT- HEURISTIK.....	64
ABB. 41: EINFLUSS DER BEGRENZUNG DER VORONOIZELLE AUF DIE VERSCHIEBUNGSVEKTOREN	65
ABB. 42: BERÜCKSICHTIGUNG VON HINTERGRUNDOBJEKTEN BEI DER VERDRÄNGUNG..	66
ABB. 43: AUFTeilUNG DER PUNKTE IN EUROPA IN UNABHÄNGIGE TEILMENGEN	67
ABB. 44 FALLUNTERSCHIEDUNGEN BEI DER SUCHE NACH PFADEN ZUR ZERLEGUNG DER PUNKTMENGE	69
ABB. 45: BEGRENZUNG DER VERSCHIEBUNGSDISTANZ AUF DEN SIGNATURRADIUS	70
ABB. 46: VERGRÖßERUNG DER ÜBERLAPPUNGSFLÄCHE VON SIGNATUREN DURCH VERKLEINERUNG DER ÜBERLAPPUNGSFLÄCHE IHRER KONSERVATIVEN APPROXIMATIONEN	71
ABB. 47: BERECHNUNG DER VERSCHIEBUNGSVEKTOREN FÜR UNTERSCHIEDLICH GROßE SIGNATUREN MITTELS DES VORONOI-DIAGRAMMS	72
ABB. 48: VERGLEICH DER VERDRÄNGUNG UNTERSCHIEDLICH GROßER SIGNATUREN MITTELS POWER-DIAGRAMM (MITTE) UND VORONOI-DIAGRAMM (LINKS)	73
ABB. 49: POSITIONIERUNGEN DER HALBKREIS-SIGNATUR INNERHALB EINER RECHTECKIGEN ZELLE	74
ABB. 50: BERECHNUNG DER DISTANZ ZUR VERSCHIEBUNG DER KANTEN DER ZELLGEOMETRIE	74
ABB. 51: VERDRÄNGUNG ELLIPTISCHER SYMBOLE. 1) EINGABE. 2) TRANSFORMIERTER RAUM. 3) VERDRÄNGTE POIs. 4) RÜCKTRANSFORMIERTER RAUM.....	75
ABB. 52: ANORDNUNGEN UND SIGNATUREN FÜR DIE SORTIERAUFGABEN DER UMFRAGE	78
ABB. 53: VERTEILUNG DER POIs AUS DENEN DIE TESTDATEN SELEKTIERT WERDEN	81
ABB. 54: TESTDATENSÄTZE FÜR FLUGHÄFEN (LINKS) UND GIPFEL (RECHTS).....	82
ABB. 55: ZERLEGUNG DER DATENSÄTZE IN UNABHÄNGIGE CLUSTER	83
ABB. 56 ANZAHL DER DURCHGEFÜHRTEN ITERATIONEN PRO POI	84
ABB. 57: ZERLEGUNG DER DATENSÄTZE IN UNABHÄNGIGE CLUSTER NACH DER ENTFERNUNG DER AN UNLÖSBAREN ÜBERLAPPUNGEN BETEILIGTEN POIs.....	85
ABB. 58: WAHL DES GEBIETES ZUR BERECHNUNG DER HI-WERTE.....	88
ABB. 59 HI-WERTE DER DATENSÄTZE VOR DER VERDRÄNGUNG	88

ABB. 60: VERGLEICH DER ÄNDERUNG DER HI-WERTE BEI VERWENDUNG DER GIK- UND SCHWERPUNKT-HEURISTIK	89
ABB. 61: VERTEILUNG DER GIPFELSIGNATUREN VOR UND NACH DER VERDRÄNGUNG....	89
ABB. 62: KERNEL DENSITY MAPS: FLUGHÄFEN/GIPFEL VOR (A/C) UND NACH (B/D) DER VERDRÄNGUNG	90
ABB. 63: FLUGHÄFEN VOR (LINKS) UND NACH DER VERDRÄNGUNG (RECHTS)	91
ABB. 64: STÄDTE VOR DER VERDRÄNGUNG	92
ABB. 65: STÄDTE NACH DER VERDRÄNGUNG	93
ABB. 66: NORMIERTE WERTEVERTEILUNG DER GIPFELHÖHEN (LINKS) UND EINWOHNERZAHLEN (RECHTS).....	94
ABB. 67: UNTERSCHIEDLICHE GROßE STADTSIGNATUREN VOR DER VERDRÄNGUNG.....	95
ABB. 68: UNTERSCHIEDLICH GROßE STADTSIGNATUREN NACH DER VERDRÄNGUNG MITTELS VORONOI-DIAGRAMM.....	96
ABB. 69: UNTERSCHIEDLICH GROßE STADTSIGNATUREN NACH DER VERDRÄNGUNG MITTELS POWER-DIAGRAMM	96
ABB. 70: UNTERSCHIEDLICH GROßE GIPFELSIGNATUREN VOR (MITTE), NACH DER VERDRÄNGUNG MITTELS POWER-DIAGRAMM OHNE (LINKS) UND MIT ANPASSUNG DER ZELLGEOMETRIE (RECHTS).....	98
ABB. 71: HÜGELSIGNATUREN MIT RADIUS R VOR (LINKS) UND NACH DER VERDRÄNGUNG (RECHTS).....	100
ABB. 72: HÜGEL MIT RADIUS R_H VOR (LINKS) UND NACH DER VERDRÄNGUNG (RECHTS)	100
ABB. 73: HÜGEL MIT RADIUS R_V VOR (LINKS) UND NACH DER VERDRÄNGUNG (RECHTS)	101

Literaturverzeichnis

- [AL10] Abascal-Mena, R., López-Ornelas, E. 2010. Geo information extraction and processing from travel narratives. In Proceedings of the International Conference on Electronic Publishing 14, pp. 363-373.
- [AO02] Ai, T., Van Oosterom, P. 2002. Displacement methods based on field analysis. In Commission on map generalization, Joint Workshop on multi-scale representations of spatial data, Ottawa, July 7-8, 2002.
- [Au87] Aurenhammer, F. 1987. Power Diagrams: Properties, Algorithms and Applications. SIAM. Journal on Computing 16(1), pp. 78-96.
- [Au91] Aurenhammer, F. 1991. Voronoi diagrams—a survey of a fundamental geometric data structure. ACM Computing Surveys (CSUR), 23(3), (pp. 345-405).
- [AW11] Abdalla, A., Weiser, P. 2011. Towards emotional and opinion-based layers in city GIS. In Proceedings of the International Conference on Urban Planning, Regional Development and Information Society 16, pp. 1039-1046.
- [BBC+04] De Berg, M., Bose, P., Cheong, O., Morin, P. 2004. On simplifying dot maps. Computational Geometry, 27(1), pp. 43-62.
- [BCM+12] Benamara, F., Chardon, B., Mathieu, Y., Popescu, V., Asher, N. 2012. How do negation and modality impact on opinions? In Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics. ExProM '12. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 10-18.
- [Bel11] Bertin, J. 2011. Semiology of graphics: diagrams networks maps, Esri Press. Redlands.
- [BG95] Bailey, T.C., Gatrell, A.C., 1995. Interactive Spatial Data Analysis. Reading, Massachusetts, Addison-Wesley.
- [BKCM+13] Borgo, R., Kehrer, J., Chung, D. H., Maguire, E., Laramée, R. S., Hauser, H., Ward, M., Chen, M. 2013. Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. In Eurographics (STARs) pp. 39-63.
- [BKO+00] De Berg, M., Van Kreveld, M., Overmars, M., Schwarzkopf, O. C. 2000. Computational geometry. In Computational geometry, Springer, Berlin, Heidelberg, pp. 1-17.
- [BL99] Bradley, M., Lang, P. 1999. Affective norms for English words (anew): Stimuli, instruction manual and affective ratings. Technical report c-1, Gainesville, FL: University of Florida.

- [BLD+16] Budninskiy, M., Liu, B., De Goes, F., Tong, Y., Alliez, P., Desbrun, M. 2016. Optimal voronoi tessellations with hessian-based anisotropy. *ACM Transactions on Graphics (TOG)*.
- [BM97] Burghardt, D., Meier, S. 1997. Cartographic displacement using the snakes concept. In *Proceedings of the Semantic modeling for the acquisition of topographic information from images and maps*. Birkhauser Verlag, (pp. 59-71).
- [BPE04] Burghardt, D., Purves, R. S., Edwardes, A. J. 2004. Techniques for on the-fly generalisation of thematic point data using hierarchical data structures. In *Proceedings of the GIS Research UK 12th Annual Conference*, pp. 28-30.
- [Br92] Brill, E. 1992. A simple rule-based part-of-speech tagger. In *Proceedings of the 3rd Conference on Applied Natural Language Processing. ANLP-92*, pp. 152-155.
- [BRW17] Boissonnat, J. D., Rouxel-Labbé, M., Wintraecken, M. 2017. Anisotropic triangulations via discrete Riemannian Voronoi diagrams. *arXiv preprint arXiv:1703.06487*.
- [BTK15] Berrios, R., Totterdell, P., Kellett, S. 2015. Eliciting mixed emotions: a meta-analysis comparing models, types, and measures. *Frontiers in psychology*, 6.
- [Bu02] Burke, R. 2002. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, 12(4), pp. 331-370.
- [BW10] Bereuter, P., Weibel, R., 2010. Generalisation of point data for mobile devices: a problem-oriented approach. *13th Workshop on Progress in Generalisation and Multiple Representation*, Zurich, 12-13 September, pp. 1-8.
- [BW11] Bereuter, P., Weibel, R. 2011. A diagnostic toolbox for assessing point data generalisation algorithms. In *Proc. 25th Int. Cartographic Conference*, Paris, 3-8 July 2011.
- [BW13a] Bereuter, P., Weibel, R. 2013. Real-time Generalization of Point Data in Mobile and Web Mapping Using Quadrees. *Cartographic and Geographic Information Science* 40(4), pp. 271-281.
- [BW13b] Bereuter, P., Weibel, R., 2013. Assessing the cartographic performance of real-time quadtree-based generalisation of point data. *ICA Workshop on Generalisation and Map Production*, Dresden, Germany.
- [BZP+11] Bergner, B. S., Zeile, P., Papastefanou, G., Rech, W., Streich, B. 2011. Emotional barrier-GIS – a new approach to integrate barrier-free planning in urban planning processes. In: M. Schrenk, V.V. Popovich, P. Zeile, eds. *Proceedings of REAL CORP 2011*, 18–20 May. Schwechat: CORP – Competence Center of Urban and Regional Planning, pp. 247–257.

-
- [CDG+14] Cresci, S., D’Errico, A., Gazzé, D., Duca, A. L., Marchetti, A., Tesconi, M. 2014. Tourpedia: a web application for sentiment visualization in tourism domain. In *Proceedings of The OpeNER Workshop in The 9th edition of the Language Resources and Evaluation Conference (LREC 2014)*, pp. 18-21.
- [Ch84] Christ, R.E. 1984. Research for evaluating visual display codes: an emphasis on colour coding. *Information design: The design and evaluation of signs and printed material*. John Wiley & Sons, Hoboken, NJ, (pp. 209-228).
- [CGG+15] Cervantes, O., Gutiérrez, F., Gutiérrez, E., Castillo, E., J. Sánchez, J.A., Wan, W. 2015. Expression: Visualizing Affective Content from Social Streams. *Proceedings of the Latin American Conference on Human Computer Interaction (CLIHIC)*.
- [CLS+12] Cao, N., Lin, Y., Sun, X., Lazer, D., Liu, S., Qu, H. 2012. Whisper: Tracing the Spatiotemporal Process of Information Diffusion in Real Time. *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, issue 12, pp. 2649-2658.
- [CM95] Cleveland, W.S., McGill, R. 1995. Graphical perception and graphical methods for analyzing scientific data. *Science*, 229, pp. 828-833.
- [CS17] Chaidee, S., Sugihara, K. 2017. Approximation of fruit skin patterns using spherical Voronoi diagrams. *Pattern Analysis and Applications*, 20(3), pp. 783-795.
- [CSS+14] Cornelia Caragea, Anna Cinzia Squicciarini, Sam Stehle, Kishore Neppalli, Andrea H. Tapia. 2014. Mapping Moods: Geo-Mapped Sentiment Analysis During Hurricane Sandy. *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.
- [CSW99] Chin, F., Snoeyink, J., Wang, C. A. 1999. Finding the medial axis of a simple polygon in linear time. *Discrete & Computational Geometry*, 21(3), pp. 405-420.
- [CWV+17] Cho, I., Wesslen, R., Volkova, S., Ribarsky, B., Dou, W. 2017. Crystal-Ball: A Visual Analytic System for Future Event Discovery and Analysis from Social Media Data. *Proceedings of the IEEE Conference on Visual Analytics Science and Technology (VAST)*.
- [De34] Delaunay, B. 1934. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk*, 7(1-2), pp. 793-800.
- [De07] Dergiades, N. 2007. The Soddy circles. *Forum Geometricorum* 7, pp. 191-197.
- [DFG99] Du, Q., Faber, V., Gunzburger, M. 1999. Centroidal Voronoi tessellations: applications and algorithms. *SIAM review*, 41(4), pp. 637-676.

- [DK10] Desrosiers, C., Karypis, G. 2011. A comprehensive survey of neighborhood-based recommendation methods. In *Recommender systems handbook*. Springer, Boston, MA, pp. 107-144.
- [DPI6] Dai, X., Prout, R. 2016. Unlocking Super Bowl Insights: Weighted Word Embeddings for Twitter Sentiment Classification. *Proceedings of the Multidisciplinary International Social Networks Conference, SocialInformatics, Data Science (MISNC, SI, DS)*.
- [DS15] Deza, M., Sikirić, M. D. 2015. Voronoi polytopes for polyhedral norms on lattices. *Discrete Applied Mathematics*, 197, pp. 42-52.
- [DSH+17] Dai, Y., Shao, J., Hu, G., Guo, L. 2017. A resource-aware approach for authenticating privacy preserving GNN queries. *World Wide Web*, pp. 1-18.
- [DTM14] Du, S., Tao, Y., Martinez, A. M. 2014. Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences*, 111(15), pp. 1454-1462.
- [EBW05] Edwardes, A., Burghardt, D., Weibel, R. 2005. Portrayal and generalisation of point maps for mobile information services. In *Map-based Mobile Services*. Springer, Berlin, Heidelberg. pp. 11-30.
- [EF71] Ekman, P., Friesen, W. V. 1971. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), p. 124.
- [EJR08] Emelianenko, M., Ju, L., Rand, A., 2008. Nondegeneracy and weak global convergence of the Lloyd algorithm in R^D . *SIAM J. Numer. Anal.* 46(3), pp. 1423-1441.
- [ES06] Esuli, A., Sebastiani, F. 2006. SENTIWORDNET: A publicly available lexical resource for opinion mining. In *Proceedings of Language Resources and Evaluation*. European Language Resources Association.
- [FF+10] Felfernig, A., Friedrich, G., Jannach, D., Zanker, M. 2011. Developing constraint-based recommenders. In *Recommender systems handbook*. Springer US, pp. 187-215.
- [FI64] Fleischer, P., 1964. Sufficient conditions for achieving minimum distortion in a quantizer. *IEEE International Convention record*, 1, IEEE, Washington, D.C., pp. 104-111.
- [Fo73] Forney, G.D. 1973. The Viterbi algorithm. In *Proceedings of the IEEE*. 61, 3 (Mar. 1973). pp. 268-278.
- [Fo87] Fortune, S. 1987. A sweepline algorithm for Voronoi diagrams. *Algorithmica*, 2(1-4), p. 153.
- [Fo99] Fodor, F. 1999. The densest packing of 19 congruent circles in a circle. *Geometriae Dedicata*, 74(2), pp. 139-145.

-
- [Fo00] Fodor, F. 2000. The Densest Packing of 12 Congruent Circles in a Circle. *Contributions to Algebra and Geometry*, 41(2), pp. 401-409.
- [Fo03] Fodor, F. 2003. The densest packing of 13 congruent circles in a circle. *Contributions to Algebra and Geometry*, 44(2), pp. 431-440.
- [FS04] Fuchs, G., Schumann, H. 2004. Visualizing abstract data on maps. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*. IEEE. pp. 139-144.
- [FSR+07] Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., Ellsworth, P. C. 2007. The world of emotion is not two-dimensional. *Psychological Science* 18,12. pp. 1050-1057.
- [FZC+17] Fu, S., Zhao, J., Cui, W., Qu, H. 2017. Visual Analysis of MOOC Forums with iForum. *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 201-210.
- [GM95] Guptill, C., Morrison, J.L. 1995. *Elements of spatial data quality*. Elsevier Science Ltd.
- [Go07] Goodchild, M. F. 2007. Citizens as voluntary sensors: spatial data infrastructure in the world of Web 2.0. *IJSDIR* 2, pp. 24–32.
- [Ha99] Harrie, L. 1999. The constraint method for solving spatial conflicts in cartographic generalization. *Cartography and Geographic Information Science*. 26, pp. 55-69.
- [Ha03] Harrie, L. 2003. Weight-setting and quality assessment in simultaneous graphic generalization. *The cartographic Journal* 40 (3), pp. 221-233.
- [HG12] Huang, H., Gartner, G. 2012. A technical survey on decluttering of icons in online map-based mashups. In *Online Maps with APIs and WebServices*. Springer, Berlin, Heidelberg. pp. 157-175.
- [HGM02] Hake, G., Grünreich, D., Meng, L. 2002. *Kartographie*. Berlin/New York.
- [HHM+16] Hoeber, O., Hoeber, L., El Meseery, M., Odoh, K., Gopi, R. 2016. Visual Twitter Analytics (Vista): Temporally changing sentiment and the discovery of emergent themes within sport event tweets. *Online Information Review*, vol. 40, no. 1, pp. 25-41.
- [Hø00] Højholt, P. 2000. Solving Space Conflicts in Map Generalization: Using a Finite Element Method. *Cartography and Geographic Information Science*, 27, 1, pp. 65-73.
- [HR]+13] Hao, M.C., Rohrdantz, C., Janetzko, H., Keim, D.A., Dayal, U., Haug L., Hsu, M., Stoffel, F.. 2013. Visual sentiment analysis of customer feedback streams using geo-temporal term associations. *Information Visualization*, vol. 12, issue 3-4, pp. 273-290.

- [HS10] Harrie, L., Stigmar, H. 2010. An evaluation of measures for quantifying map information. *Journal of Photogrammetry and Remote Sensing*, 65, pp. 266-274.
- [HSD15] Harrie, L., Stigmar, H., Djordjevic, M. 2015. Analytical estimation of map readability. *ISPRS Int. J. Geo-Inf.* 4, pp. 418-446.
- [HSM+17] Hundt, M., Schneider, B., El-Assady, M., Keim, D., Diehl, A. 2017. Visual Analysis of Geolocated Echo Chambers in Social Media. Poster Abstracts of the Eurographics Conference on Visualization (EuroVis).
- [HW07] Harrie, L., Weibel, R. 2007. Modelling the overall process of generalisation, Generalisation of geographic information: Cartographic Modelling and applications, Elsevier, pp 67-88.
- [HZO+97] Hupka, R. B., Zaleski, Z., Otto, J., Reidl, L., Tarabrina, N. V. 1997. The colors of anger, envy, fear, and jealousy: A cross-cultural study. *Journal of cross-cultural psychology*, 28(2), pp. 156-171.
- [IHH10] Iosifescu-Enescu, I., Hugentobler, M., Hurni, L. 2010. Web cartography with open standards – A solution to cartographic challenges of environmental management. *Environmental Modelling & Software* 25(9), pp. 988-999.
- [Im96] Imai, T. 1996. A topology oriented algorithm for the Voronoi diagram of polygons. In *Proceedings of the 8th Canadian Conference on Computational Geometry*. Carleton University Press, pp. 107-112.
- [Iz77] Izard, C. E. 2013. Human emotions. Springer Science & Business Media.
- [Jo98] Joao, E. 1998. Causes and consequences of map generalization, Taylor & Francis.
- [Ka93] Kaplan, R. 1993. The role of nature in the context of the workplace. *Landscape and Urban Planning*, 26, pp. 193-201.
- [KA13] Korpi, J., Ahonen-Rainio, P. 2013. Clutter reduction methods for point symbols in map mashups. *The Cartographic Journal*, 50(3), pp. 257-265.
- [KFK+08] Kawai, Y., Fujita, Y., Kumamoto, T., Zhang, J., Tanaka, K. 2008. Using a sentiment map for visualizing credibility of news sites on the web. In *Proceedings of the ACM workshop on Information credibility on the Web*. ACM, New York, NY 2, pp. 53-58.
- [KG12] Klettner S., Gartner, G. 2012. Modeling affective responses to space. In *Proceedings of the International Conference on Urban Planning, Regional Development and Information Society*. 17, pp. 485-491.
- [KHA14] Korpi, J., Haybatollahi, M., Ahonen-Rainio, P. 2014. Identification of partially occluded map symbols. *Cartographic Perspectives*, 0(76): doi:10.14714/CP76.59, pp. 19-32.

-
- [KMM+04] Kamps, J., Marx, M., Mokken, R. J., Rijke, M. D. 2004. Using wordnet to measure semantic orientation of adjectives. In Proceedings of the LREC'04. (Lisboa, Portugal). pp. 1115-1118.
- [KPI0] Kelso, N. V., Patterson, T. 2010. Introducing natural earth data-naturalearthdata. com. *Geographia Technica*, 5, pp. 82-89.
- [KPK17] Kucher, K., Paradis, C., Kerren, A. 2017. The state of the art in sentiment visualization. In *Computer Graphics Forum*.
- [KTL+14] Kuang, C., Tang, J., Liu, Z., Sun, M. 2014. ImgWordle: image and text visualization for events in microblogging services. *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI)*, pp. 371-372.
- [KVH+95] Karlsson, F., Voutilainen, A., Heikkilä, J., Anttila, A. 1995. Constraint Grammar: a language-independent system for parsing running text. *Natural Language Processing*. 4. Mouton de Gruyter, Berlin and New York.
- [LD81] Lee, D. T., Drysdale, III, R. L. 1981. Generalization of Voronoi diagrams in the plane. *SIAM Journal on Computing*, 10(1), pp. 73-87.
- [Le04] Levine, N., 2004. CrimeStat III: A spatial statistics program for the analysis of crime incident locations. Houston, Ned Levine and associates, Wahington DC, National Institute of Justice.
- [LGS10] Lops, P., De Gemmis, M., Semeraro, G. 2011. Content-based recommender systems: State of the art and trends. In *Recommender systems handbook*. Springer, Boston, MA, pp. 73-105.
- [LH02] Li, Z., Huang, P., 2002. Quantitative measures for spatial information of maps. *International Journal of Geographical Information Science*, 16(7), pp. 699-709.
- [LHW+15] Lu, Y., Hu, X., Wang, F., Kumar, S., Liu, H., Maciejewski, R. 2015. Visualizing Social Media Sentiment in Disaster Scenarios. *Proceedings of the International Conference on World Wide Web (WWW)*, pp. 1211-1215.
- [Li12] Liu, B. 2012. *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers, 2012.
- [LJ01] Lonergan, M.E., Jones, C.B. 2001. An iterative displacement method for conflict resolution in map generalization. *Algorithmica*, 30(2) (pp. 287-301)
- [LI82] Lloyd, S. 1982. Least square quantization in PCM. *IEEE Trans. Inform. Theory*, 28, pp. 129-137.
- [LSB+16] Lu, Y., Steptoe, M., Burke, S., Wang, H., Tsai, J., Davulcu, H., Montgomery, D., Corman, S.r., Maciejewski, R. 2016. Exploring Evolving Media Discourse Through Event Cueing. *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 220-229.

- [LWW+16] Li, Q., Wu, Y., Wang, S., Lin, M., Feng, X., Wang, H. 2016. VisTravel: visualizing tourism network opinion from the user generated content. *Journal of Visualization*, vol. 19, no. 3, pp. 489–502.
- [LYH+06] Lin, C.-H., Yan, C.-R., Hsu, J.-H., Lee, T.-Y. 2006. Multiresolution Remeshing Using Weighted Centroidal Voronoi Diagram. In *Proc. of Int. Conf. on Computational Science*, pp. 295–301.
- [Ma86] MacKinlay, J.D. 1986. Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*. 5(2), pp. 110-141.
- [Ma94] Mackaness, W.A. 1994. An algorithm for conflict identification and feature displacement in automated map generalization. *Cartography and Geographic Information Systems*, 21(4), pp. 219-232.
- [MBB+11] Marcus, A., Bernstein, M.S., Badar, O., Karger, D.R., Madden, S., Miller, R.C. 2011. TwitInfo: Aggregating and Visualizing Microblogs for Event Exploration. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pp. 227-236.
- [MBQ01] Matei, S., Ball-Rokeach, S.J., Qiu, J.L. 2001. Fear and misperception of Los Angeles urban space: A spatial-statistical study of communication-shaped mental maps. *Communication Research* 28(4), pp. 429-463.
- [Mc06] McCloud, S. 2006. *Making Comics: Storytelling Secrets of Comics, Manga and Graphic Novels*. William Morrow Paperbacks.
- [Me94] Melissen, H. 1994. Densest packings of eleven congruent circles in a circle. *Geometriae Dedicata*, 50(1), pp. 15-25.
- [Me16] Sebastian, M. 2016. The Marker Cluster: A Critical Analysis and a New Approach to a Common Web-based Cartographic Interface Pattern. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, 7(1), pp. 28-43.
- [MFB17] Melo, S. N. D., Frank, R., Brantingham, P. 2017. Voronoi Diagrams and Spatial Analysis of Crime. *The Professional Geographer*, 69(4), pp. 579-590.
- [MHR00] Madden, T. J., Hewett, K., Roth, M. S. 2000. Managing images in different cultures: A cross-national study of color meanings and preferences. *Journal of international marketing*, 8(4), pp. 90-107.
- [MP01] Mackaness, W.A., Purves, R. 2001. Automated displacement for large numbers of discrete map objects. *Algorithmica*. 30, pp. 302-311.
- [MS91] Marks, J., Shieber, S. M. 1991. The computational complexity of cartographic label placement. Technical Report TR-05-91, Center for Research in Computing Technology, Harvard University.
- [MS14] Moacdieh, N., Sarter, N. 2015. Display clutter: A review of definitions and measurement techniques. *Human factors*, 57(1), pp. 61-100.

-
- [MSK+17] Martins, R.M., Simaki V., Kucher, K., Paradis, C., Kerren, A. 2017. StanceXplore: Visualization for the Interactive Exploration of Stance in Social Media. Proceedings of the Workshop on Visualization for the Digital Humanities (VIS4DH).
- [MSV15] Mahdian, M., Schrijvers, O., Vassilvitskii, S. 2015. Algorithmic cartography: Placing points of interest and ads on maps. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, pp. 755-764.
- [MWK09] Mody, R. N., Willis, K. S., Kerstein, R. 2009. WiMo: location-based emotion tagging. In Proceedings of the 8th international Conference on Mobile and Ubiquitous Multimedia. ACM, p. 14.
- [Ne82] Newman, D. 1982. The hexagon theorem. IEEE Transactions on information theory, 28(2), pp. 137-139.
- [NE04] NAz, K. A. Y. A., Epps, H. 2004. Relationship between color and emotion: A study of college students. College Student J, 38(3), p. 396.
- [NGH+13] Nummenmaa, L., Glerean, E., Hari, R., Hietanen, J. K. 2014. Bodily maps of emotions. Proceedings of the National Academy of Sciences, 111(2), pp. 646-651.
- [Ni11] Nielsen, F. Å. 2011. A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. ArXiv e-prints -arXiv: 1103.2903.
- [No09] Nold, C. 2009. Emotional cartography - technologies of the self. [online] Available through: <http://emotionalcartography.net/> [Accessed 17 June 2014].
- [NSH02] Nowell, L., Schulman, R., Hix, D. 2002. Graphical encoding for information visualization: An empirical study. In Proceedings of the IEEE Symposium on Information Visualization. pp. 43-51.
- [OCO+82] O'Rourke, J., Chien, C. B., Olson, T., Naddor, D. 1982. A new linear algorithm for intersecting convex polygons. Computer Graphics and Image Processing, 19(4), pp. 384-391.
- [OGC02] Open Geospatial Consortium, 2002. ISO6-042 OpenGIS Web Map Service (WMS) Implementation Specification, Version 1.1.3. Open Geospatial Consortium [online] Available through: Open Geospatial Consortium <http://www.opengeospatial.org/specs/> [Accessed 15 August 2012]
- [OH11] Ortag, F., Hoasheng, H. 2011. Location-based emotions relevant for pedestrian navigation. In Proceedings of the 25th international cartographic conference, Paris.
- [Ok92] Okabe, A. 1992. Spatial tessellations. John Wiley & Sons, Ltd.

-
- [OMM75] Osgood, C. E., May, W. H., Miron, M. S. 1975. Cross-cultural universals in affective meaning. Urbana, University of Illinois Press.
- [OSL+10] Ostlaender, N.; Smith, R.S.; De Longueville, B.; Smits, P.; 2010. What Volunteered Geographic Information is (good for) - designing a methodology for comparative analysis of existing applications to classify VGI and its uses, Geoscience and Remote Sensing Symposium (IGARSS), IEEE International, 25-30 July 2010, pp.1422-1425.
- [OU03] O'Sullivan, D., Unwin, D.J., 2003. Geographic Information Analysis. John Wiley and Sons, Hoboken, NJ.
- [PAH+09] Podolskaya, E.S., Anders, K.-H., Haunert, J.-H., Sester, M. 2009. Quality Assessment for polygon generalization. The international Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34.
- [Pi69] Pirl, U. 1969. Der Mindestabstand von n in der Einheitskreisscheibe gelegenen Punkten. Mathematische Nachrichten, 40(1-3), pp. 111-124.
- [PKS16] Pino, C., Kavasidis, I., Spampinato, C. 2016. GeoSentiment: a Tool for Analyzing Geographically Distributed Event-related Sentiments. Proceedings of the IEEE Annual Consumer Communications Networking Conference (CCNC), pp. 270-271.
- [PI01] Plutchik, R. 2001. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American scientist, 89(4), pp. 344-350.
- [PLV02] Pang, B., Lee, L., Vaithyanathan, S. 2002. Thumbs up?: sentiment classification using machine learning techniques. In Proceedings of the ACL-02 conference on Empirical methods in natural language processing Association for Computational Linguistics, 10, pp. 79-86.
- [PT12] Paltoglou, G., Thelwall, M. 2012. Twitter, MySpace, Digg: Unsupervised sentiment analysis in social media. ACM Trans. Intell. Syst. Technol. 3, 4 (Sep. 2012). 66, pp. 1-19.
- [QRQ15] Qiu, R.G., Ravi, R.R., Qiu, L.L. 2015. Aggregating and visualizing public opinions and sentiment trends on the US higher education. Proceedings of the International Conference on Information Integration and Web-based Applications & Services (iiWAS).
- [Ra02] Ratcliffe, J., 2002. HotSpot Detective 2.0 for MapInfo professional7.0. Philadelphia, JHR Systems.
- [RP96] Ruas, A., Plazanet, C. 1996. Strategies for automated generalization. In Proc of the 7th spatial data handling symposium, Delft, the Netherlands, pp 319-336.

-
- [RRS11] Ricci, F., Rokach, L., Shapira, B. 2011. Introduction to recommender systems handbook. In *Recommender systems handbook*. Springer US, pp. 1-35.
 - [Ru80] Russel, J. A. 1980. A circumplex model of affect. *Journal of Personality and Social Psychology*. 39, 6, pp. 1161-1178.
 - [Ru98] Ruas, A. 1998. OO-constraint modelling to automate urban generalization process. In *Proceedings of the Eighth International Symposium on Spatial Data Handling (Vancouver, Canada, July 12–15, 1998)*, pp. 225-235.
 - [Ru08] Ruas, A. 2008. Map generalization. In *Encyclopedia of GIS*. Springer US, pp. 631-632.
 - [SBB+15] Steed, C.A., Beaver, J., Bogen P.L., Drouhard, M., Pyle, J. 2015. Text Stream Trend Analysis using Multiscale Visual Analytics with Applications to Social Media Systems. *Proceedings of the IUI Workshop on Visual Text Analytics*.
 - [Se05] Sester, M. 2005. Optimization approaches for generalization and data abstraction, *International Journal of Geographical Information Science*, 19(8-9), pp. 871-897.
 - [SH75] Shamos, M. I., Hoey, D. 1975. Closest-point problems. In *Foundations of Computer Science, 1975., 16th Annual Symposium on*. IEEE, pp. 151-162.
 - [SH76] Shamos, M. I., Hoey, D. 1976. Geometric intersection problems. In *Foundations of Computer Science, 1976., 17th Annual Symposium on*. IEEE, pp. 208-215.
 - [SH08] Sui, D. Z., Holt, J. B. 2008. Visualizing and analysing public-health data using value-by-area cartograms: toward a new synthetic framework. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 43(1), pp. 3-20.
 - [SHW+12] Scharl, A., Hubmann-Haidvogel, A., Weichselbraun, A., Wohlgenannt, G., Lang, H., Sabou, M. 2012. Extraction and interactive exploration of knowledge from aggregated news and social media content. *Proceedings of the ACM SIGCHI symposium on Engineering interactive computing systems (EICS)*, pp. 163-168.
 - [SK99] Sarjakoski, T., Kilpeläinen, T. 1999. Holistic cartographic generalization by least squares adjustment for large data sets. In *Proceedings of the 19th ICA/ACI conference (Ottawa, Canada)*. pp. 1091-1098.
 - [SKD04] Sullivan, W.C., Kuo, F.E., De Pooter, S. 2004. The fruit of urban nature: Vital neighborhood spaces. *Environment & Behavior*. 36, 5, pp. 678-700.
 - [SQC16] Sijtsma, B., Qvarfordt, P., Chen, F. 2016. Tweetviz: Visualizing Tweets for Business Intelligence. *Proceedings of the International ACM SIGIR conference*.

- ence on Research and Development in Information Retrieval (SIGIR), pp. 1153-1156.
- [SW07] Steiniger, S., Weibel, R. 2007. Relations among map objects in cartographic generalization. *Cartography and Geographic Information Science*, 34(3), pp. 175-197.
- [SW12] Schmidt, M., Weiser, P. 2012. Web mapping services: development and trends. In: M.P. Peterson (ed.). *Online maps with APIs and webservices*, Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg, Germany, pp. 13-21.
- [TBP+10] Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A. 2010. Sentiment in short strength detection informal text. *J.Am. Soc. Inf. Sci. Technol.* 61, 12 (Dec. 2010), pp. 2544-2558.
- [TBT+11] Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M. 2011. Lexicon-based methods for sentiment analysis. *Comput.Linguist.* 37, 2 (Jun. 2011), pp. 267-307.
- [TL03] Turney, P. D., Littman, M. L. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Trans. Inf. Syst.* 21, 4 (Oct. 2003), pp. 315-346.
- [Tu93] Turlach, B.A., 1993. Bandwidth selection in kernel density estimation: a review. *The Annals of Statistics*, pp. 1883-1905.
- [UI84] Ulrich, R.S. 1984. View through a window may influence recovery from surgery. *Science*. 224, pp. 420-421.
- [WA02] Wojnarowska M., Ady, B.E. 2002. Interoperable solutions in web-based mapping. *Proc. ISPRS Comm. IV Symposium*, Toronto/Ottawa.
- [WD98] Weibel, R., Dutton, G. 1998. Constraint-based automated map generalization. In *Proc. Of the 8th spatial data handling symposium*. Vancouver, pp. 214-224.
- [WF94] Wassermann, S., Faust, K., 1994. *Social network analysis: methods and applications*. Cambridge Univ. Press.
- [WFL+12] Wu, J., Fu, Z., Liu, Z., Pan, J., Long, H., Lin, X., He, H., Chen, X., Tang, J. 2012. City Flow: Prototype Exploration for Visualizing Urban Traffic Conversations. *Proceedings of the International Conference on Privacy, Security, Risk and Trust (PASSAT) and the International Conference on Social Computing (SocialCom)*, pp. 481-489.
- [WFL+13] Wu, J., Fu, Z., Liu, Z., Lin, X., Tang, J., Pan, J., Zhao, C. 2013. Creating reflections in public emotion visualization: prototype exploration on traffic theme. *Proceedings of the ACM Conference on Creativity & Cognition (C&C)*, pp. 357-361.

-
- [WGS15] Weiler, A., Grossniklaus, M., Scholl, M.H. 2015. The Stor-e-Motion Visualization for Topic Evolution Tracking in Text Data Streams. *Proceedings of the International Conference on Information Visualization Theory and Applications (IVAPP)*, pp. 29-39.
- [WWL+10] Wu, Y., Wei, F., Liu, S., Au, N., Cui, W., Zhou, H., Qu, H. 2010. OpinionSeer: Interactive Visualization of Hotel Customer Feedback. *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, issue 6, pp. 1109-1118.
- [XTS08] Xiang, Z., Tinghua, A., Stoter, J. 2008. The evaluation of spatial distribution density in map generalization. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37(B2), pp. 181-187.
- [YHW+14] Yang, W., Hu, J., Wang, S., Stojmenovic, M. 2014. An alignment-free fingerprint bio-cryptosystem based on modified Voronoi neighbor structures. *Pattern Recognition*, 47(3), pp. 1309-1320.
- [YW08] Yan, H., Weibel, R. 2008. An algorithm for point cluster generalization based on the Voronoi diagram. *Computers & Geosciences*, 34(8), pp. 939-954.
- [YWC+16] Yu, Z., Wang, Z., Chen, L., Guo, B., Li, W.. 2016. Featuring, Detecting, and Visualizing Human Sentiment in Chinese Micro-Blog. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 10, no. 4.
- [Za01] Zacharias, J. 2001. Path choice and visual stimuli: Signs of human activity and architecture. *Journal of Environmental Psychology*. 21(34).
- [ZDW+12] Zhao, J., Dongm L., Wu, J., Xu, K. 2012. MoodLens: an emoticon-based sentiment analysis system for Chinese tweets. *Proceedings of the ACM SIGKDD international conference on Knowledge discovery and data mining (KDD)*, pp. 1528-1531.
- [ZKK+09] Zhang, J., Kawai Y., Kumamoto T., Tanaka, K. 2009. A Novel Visualization Method for Distinction of Web News Sentiment. *Proceedings of the International Conference on Web Information Systems Engineering (WISE)*, pp. 181-194.
- [ZLW13] Zhang, C., Liu, Y., Wang, C. 2013. Time-space Varying Visual Analysis of Micro-blog Sentiment. *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction (VINCI)*, pp. 64-71.
- [ZQL+16] Zhao, Y., Qin, B., Liu, T., Tang, D. 2016. Social sentiment sensor: a visualization system for topic detection and topic sentiment analysis on microblog. *Multimedia Tools and Applications*, vol. 75, no. 15, pp. 8843-8860.